Health Profession Opportunity Grants (HPOG) Impact Study

Technical Supplement to the Evaluation Design Report: Impact Analysis Plan

OPRE Report No. 2015-80

August 5, 2015

submitted by
Eleanor L. Harvill
Shawn Moulton
Laura R. Peck
Abt Associates, Inc.

submitted to
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

Health Profession Opportunity Grants (HPOG) Impact Study Technical Supplement to the Evaluation Design Report: Impact Analysis Plan

OPRE Report No. 2015-80

August 5, 2015

Eleanor Harvill, Shawn Moulton, Laura R. Peck, Abt Associates Inc.

Submitted to:
Hilary Forster, Contracting Officer's Technical Representative
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

Contract No. HHSP23320095624WC, Task Order HHSP23337012T

Project Director: Gretchen Locke
Abt Associates
55 Wheeler Street
Cambridge, MA 02138

The authors acknowledge and appreciate the research assistance and contributions of Jordan Misra, Hiren Nisar, Sarah Sahni and Daniel Weiss, and the review and input of Ed Bein, Stephen H. Bell, Gretchen Locke, Rob Olsen and Alan Werner.

Table of Contents

1. Introduction and Design

This is a Technical Supplement to the Evaluation Design Report for the Health Profession Opportunity Grants (HPOG) Impact Study. It provides information on the study's impact analyses, ideally with sufficient transparency that others can replicate our work either on the HPOG-Impact data or on other similar experimental data sets.

1.1 Description of Intervention

As part of the Affordable Care Act of 2010, Congress authorized funds for the Health Profession Opportunity Grants (HPOG) program "to conduct demonstration projects that provide eligible individuals with the opportunity to obtain education and training for occupations in the healthcare field that pay well and are expected to either experience labor shortages or be in high demand" [Affordable Care Act, Public Law 111-148, 124 Stat. 119, March 23, 2010, sect. 5507(a), adding sect. 2008(a) to the Social Security Act, 42 U.S.C. 1397g(a)]. In September 2010, the U.S. Department of Health and Human Services (HHS) Administration for Children and Families (ACF) awarded 32 grants to government agencies, community-based organizations, post-secondary educational institutions, and tribal-affiliated organizations to conduct these activities in 23 states to provide education and training services to Temporary Assistance for Needy Families (TANF) recipients and other low-income individuals.

The HPOG objectives stem from the career pathways framework of postsecondary education, a framework designed to address the challenge of preparing nontraditional student populations with varying ranges of assets and challenges related to employment and training. Specifically, HPOG is intended to:

- Target skills and competencies demanded by the healthcare industry.
- Support "career pathways"—clearly defined routes that allow participants to build a career, rather than simply getting training for a job, by advancing through successively higher levels of education and training, exiting into employment at multiple possible points.
- Result in employer- or industry-recognized, portable education credentials (e.g., certificates or degrees) and professional certifications and licenses (e.g., a credential awarded by a Registered Apprenticeship program).
- Combine support services with education and training to help participants overcome barriers to employment.
- Provide training at times and locations that are easily accessible to targeted populations.

The demonstration projects are intended to address two pervasive problems: the increasing shortfall in the supply of qualified healthcare professionals in the face of expanding demand, and the increasing requirement for a postsecondary education to secure a job with a living wage for families.

1.2 Logic Model

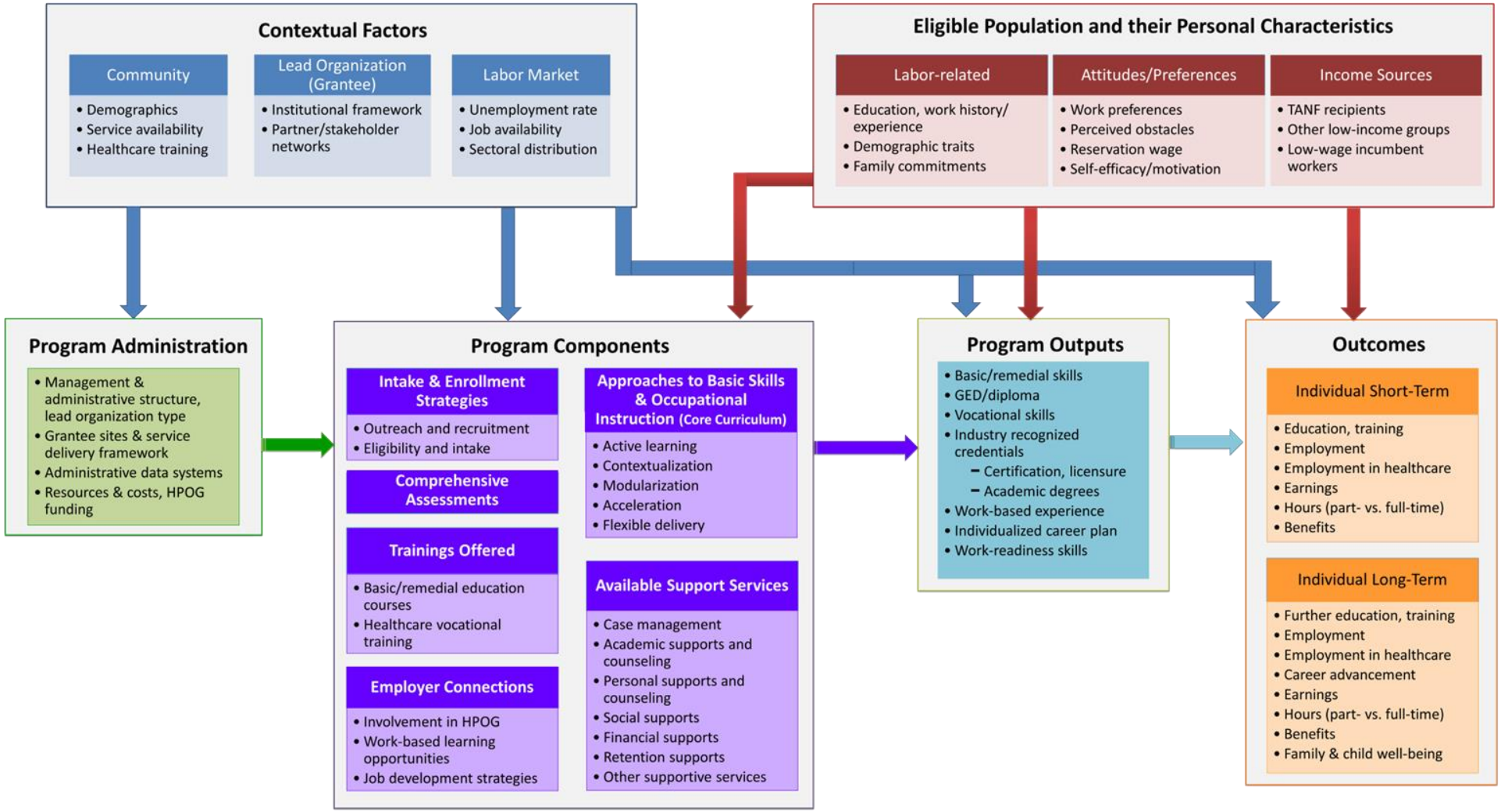
Each of the HPOG programs—and each of the administrative divisions that operate semi-autonomously within each HPOG program (see Section 1.5 for further definition)—uses its own chosen design as its way of achieving desired results for participants. The evaluation design and analysis plan capitalize on this variation in intervention approaches to learn more about what works. As detailed in Chapter 2 of the Evaluation Design Report, and repeated below as Exhibit 1.1, each program embeds its own distinctive elements in its intervention approach, making the logic model specific to that program. That is, while we have a single overarching conceptual framework, any given program offers and emphasizes selected elements of the framework. While HPOG-Impact has been able to identify some specific elements from among the program's potential causal links for direct testing through planned variation, we also have other strategies for getting answers to questions regarding the effectiveness of selected program elements. This Analysis Plan builds on the program's logic model and diverse intervention approaches to specify how the evaluation will estimate the program's overall impacts and identify dimensions on which its impacts vary.

1.3 Research Questions

This Analysis Plan describes our approach to addressing the following Research Questions:

1. What impacts do the standard HPOG programs as a group have on the outcomes of participants and their families?
2. To what extent do those impacts vary across selected subpopulations?
3. Which locally adopted program components influence average impacts?
4. To what extent does participation in a particular HPOG component (or components) change the impact on trainees?

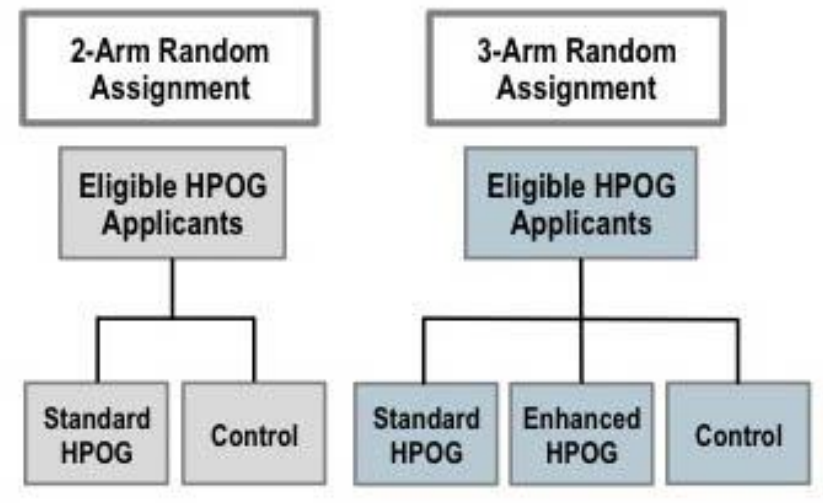
Exhibit 1.1: HPOG Career Pathways Framework Logic Model



1.4 Experimental Evaluation Design

HPOG-Impact uses an experimental evaluation design. For some grantees, eligible applicants are randomly assigned to either a treatment group that offers access to the HPOG program or to a control group that does not but can access whatever other programs and services are available in the community. For other grantees, eligible applicants are randomly assigned to three groups—the treatment and control groups, plus an enhanced treatment group of individuals offered access to the grantee's regular HPOG program supplemented with one of three components: facilitated peer support, emergency assistance, and non-cash incentives. Exhibit 1.2 shows the two-arm design that some grantees used and the three-arm design used by grantees that are testing one of the three program enhancements.

Exhibit 1.2: HPOG-Impact Evaluation Design that Uses Two- and Three-Arm Randomization



1.5 Sites Included in the Study

Twenty-three of the 32 HPOG grantees were included in HPOG-Impact.[1] Three of these grantees are also part of ACF's Pathways for Advancing Careers and Education (PACE) project, which is primarily responsible for data collection, but they are fully part of the HPOG-Impact analysis.

Although the federal funding of HPOG goes through a grantee agency, the administrative framework of the HPOG program includes additional levels that are relevant to the evaluation, as follows:

- **Grantee** – the funded unit of the national HPOG program.
- **Program** – a unique set of services, training courses and personnel. Many grantees fund and operate one program; some fund multiple programs.

[1] All of the HPOG grantees are part of some evaluation research. All of the non-tribal HPOG grantees are part of the National Implementation Evaluation (NIE), and all of the tribal grantees are part of the Evaluation of Tribal HPOG. Several grantees are also participating in research via the ACF-sponsored University Partnership Research Grants for HPOG. Those grantees collecting individual-level data as part of another evaluation were not required to participate in the HPOG-Impact Study; all the rest were required to participate.

- **Administrative division** – a set of program intake locations with a dedicated case management and/or counseling staff that advises participants, connects them to education and training services, and provides participants with support services or refers them to these services. Programs may have one or more such divisions.

We define these units here because the remainder of this Analysis Plan explains how we will exploit variation across these administrative levels to learn about the relative effectiveness of various approaches to delivering HPOG's career pathways-based training.

1.6 Sample Sizes

Exhibit 1.3 shows the samples sizes for HPOG-Impact, separately for grantees, programs, administrative divisions, and individuals. Sample sizes are also presented separately for grantees conducting two-way and three-way random assignment, and separately by whether grantees are participating in ACF's PACE project.

Exhibit 1.3: HPOG-Impact Units of Analysis and Sample Sizes

Units using Two- and Three-Arm Randomization	HPOG Only	HPOG and PACE	Total HPOG-Impact
Grantees	20	3	23
Programs	36	6	42
Administrative divisions	81	6	87
Individuals	10,694	3,117	13,811
Units using Three-Arm Randomization Only			
Grantees in three-arm experiments	10	0	10
Programs in three-arm experiments	19	0	19
Administrative divisions in three-arm experiments	31	0	31
Individuals	6,043	0	6,043

1.7 Outline of Analysis Plan

HPOG-Impact's Evaluation Design Report describes the HPOG intervention, outlines measures and data collection, and presents analytic methods for conducting the impact analysis. This document serves as a Technical Supplement to the Evaluation Design Report and specifies the impact analysis plans in full, providing key details for the operationalization and implementation of the measures and analyses introduced in the Evaluation Design Report. The document proceeds as follows: Section 2 details the study's data and outcome and baseline measures; and it also includes plans for handling missing data and documenting attrition. Section 3 describes the characteristics of the research sample and analysis of baseline data, including a report of baseline balance. Section 4 covers measures of program characteristics and participation. Section 5 first explains our approach to hypothesis testing, then elaborates on the main impact analyses—including our approach to addressing Research Questions 1 and 2—refining the details provided in the main Evaluation Design Report to create an integrated model that will serve as the basis for most subsequent analyses. Section 6 explains how we will address Research Question 3, using cross-site variation to estimate the contribution of program components (both randomly assigned to and naturally occurring) to program impacts. Section 7 discusses our analysis of individuals' participation experiences in an analysis of endogenous subgroups to address Research Question 4. A final section lists future project activities and their timeline. Appendix A details the full operationalization of measures described in the text; Appendix B details our plan for calculating attrition; Appendix C presents sample sizes and missing data rates associated with descriptive statistics; and Appendices D and E explain additional detailed elements of our planned analytic processes (variable selection, omitted variables).

2. Data and Measures

This section of the Analysis Plan identifies the main data sources, measures, response rates, and treatment of missing data in the analysis. Additional program and participation measures are detailed later in Section 4.

2.1 Data Sources

The study will collect data from a Performance Reporting System (PRS) embedded into each's grantee's program, a series of program and staff surveys, a 15-month follow-up survey, and the National Directory of New Hires (NDNH). For more information, the HPOG-Impact Evaluation Design Report elaborates on these data sources in Chapter 3 (3.1 for Baseline Measures, 3.2 for Program Measures, and 3.3 for Outcomes).

2.2 Measures

This section describes baseline and outcome measures required to estimate the impacts of HPOG and the three enhancements.

2.2.1 Baseline Measures

Individual baseline characteristics serve several purposes in the impact analysis.

- Baseline *descriptors* are constructed to provide contextual information on the individuals in the sample. For readers familiar with HPOG Annual Reports,[2] these measures retain the same specification as those earlier publications.
- A subset of baseline descriptors, *subgroup identifiers* define exogenous subgroups for moderator analyses.
- Baseline *covariates* are constructed for inclusion in impact models. The selection of baseline covariates for inclusion in analytic models is theory-driven, based on the expected relationships between the outcomes of interest and baseline measures. These baseline covariates include *pre-intervention measures of outcomes* when possible to facilitate baseline balance testing.

The next set of exhibits presents the baseline measures, organized by domain, with the designated purpose(s) and data source(s) indicated. Exhibit 2.1 includes constructs that fall within the domain of demographic measures; Exhibit 2.2 includes measures of educational background; Exhibit 2.3 describes economic status measures; Exhibit 2.4 describes measures of employment and expected time use; and Exhibit 2.5 describes life challenges-related measures. Generally, for HPOG-only program sample members, these data come from the PRS. For HPOG/PACE sample members, these data were collected by the PACE evaluation via the Basic Information Form (BIF) and Self-Administered Questionnaire (SAQ). These instruments are aligned by design, but there are some differences for specific measures. We note any differences in measure availability in the variable description column. In addition, Appendix A contains details of the operationalization of these variables, including information on any difference in operationalization between the HPOG-Impact and HPOG/PACE programs. Subsequent sections describe how these variables enter the analysis.

[2] The HPOG Annual Reports (Year Two and Three are released, Year Four is in production) provide an overview of the HPOG Program up to that point, including characteristics of participants and the activities in which they engaged, participant training and employment outcomes, and examples of grantee program implementation.

Exhibit 2.1: Baseline Demographic Measures

Construct	Variable Description	Purpose(s)	Data Source(s)
Sex	Respondent is male (binary)	Descriptor; Covariate	HPOG-Impact site: PRS HPOG/PACE site: BIF
Marital Status	Marital status (categorical)	Descriptor	HPOG-Impact site: PRS HPOG/PACE site: BIF
Dependent Children	Number of dependent children (categorical)	Descriptor	HPOG-Impact site: PRS HPOG/PACE site: BIF
Dependent Children	Parent to one or more dependent children (binary)	Covariate; Subgroup Identifier	HPOG-Impact site: PRS HPOG/PACE site: BIF
Race/Ethnicity	Ethnicity (Hispanic/Latino) and race (series of mutually exclusive binary variables)	Descriptor; Covariate; Subgroup Identifier	HPOG-Impact site: PRS HPOG/PACE site: BIF
Age	Age	Descriptor (categorical); Covariate (continuous)	HPOG-Impact site: PRS HPOG/PACE site: BIF
Age	Typical postsecondary age (younger than 25 years old at random assignment) (binary)	Subgroup Identifier	HPOG-Impact site: PRS HPOG/PACE site: BIF
Born Outside U.S.	Born outside the U.S. (binary)	Descriptor; Covariate	HPOG-Impact site: PRS HPOG/PACE site: BIF

Exhibit 2.2: Baseline Education Background Measures

Construct	Variable Description	Purpose(s)	Data Source(s)
Education	Completed education measured in years (categorical)	Descriptor	HPOG-Impact site: PRS HPOG/PACE site: BIF
Education	Educational attainment measured by degrees and credentials completed (series of mutually exclusive binary variables)	Subgroup Identifier	HPOG-Impact site: PRS HPOG/PACE site: BIF
Education	Attained postsecondary degree prior to random assignment (binary)	Descriptor; Pre-intervention measure of outcome Covariate	HPOG-Impact site: PRS HPOG/PACE site: BIF
Education	Occupational skills license or certification prior to random assignment (binary)	Descriptor; Pre-intervention measure of outcome; Covariate	HPOG-Impact site: PRS HPOG/PACE site: BIF
Education	Completed license, certification or degree prior to random assignment (binary)	Pre-intervention measure of outcome Covariate	HPOG-Impact site: PRS HPOG/PACE site: BIF
Basic Skills	Literacy assessed at 8th grade level or higher (binary) Note: This measure is not available for PACE programs.	Descriptor; Covariate	HPOG-Impact site: PRS
Basic Skills	Numeracy assessed at 8th grade level or higher (binary) Note: This measure is not available for PACE programs.	Descriptor; Covariate	HPOG-Impact site: PRS
Skills course attendance	Course attendance in - Adult Basic Education classes - English as a Second Language classes - Vocational, Technical or Trade School classes - Classes in how to succeed in school - Classes in how to succeed at work (series of binary variables)	Descriptor; Covariate	HPOG-Impact site: PRS HPOG/PACE site: BIF

Construct	Variable Description	Purpose(s)	Data Source(s)
Educational Expectations	Level of education expected to complete, measured in degrees and credentials (categorical measure)	Descriptor	HPOG-Impact site: PRS HPOG/PACE site: BIF

Exhibit 2.3: Baseline Income and Benefits Measures

Construct	Variable Description	Purpose(s)	Data Source(s)
Earnings	Average quarterly wage received during the four quarters prior to the quarter of random assignment (continuous)	Pre-intervention measure of outcome; Covariate	NDNH
Income	Household income over 12 months prior to random assignment (categorical)	Descriptor	HPOG-Impact site: PRS HPOG/PACE site: BIF
Income	Individual income over 12 months prior to random assignment (categorical) Note: This measure is not available for PACE programs.	Descriptor	HPOG-Impact site: PRS
Public Assistance	Public assistance use, by source - Welfare[a] - WIC/SNAP (series of binary variables)	Descriptor; Pre-intervention measure of outcome; Covariate; Subgroup Identifier	HPOG-Impact site: PRS HPOG/PACE site: BIF

Notes: The baseline measure of Earnings comes from NDNH data.
[a]HPOG/PACE participants reported whether they received any form of welfare, whereas HPOG/Impact participants reported only TANF.

Exhibit 2.4: Baseline Employment and Expected Time Use Measures

Construct	Variable Description	Purpose(s)	Data Source(s)
Employment	Proportion of quarters employed during the four quarters prior to the quarter of random assignment (continuous)	Descriptor; Pre-intervention measure of outcome; Covariate	NDNH
Employment	Currently employed (self-report) (binary)	Subgroup Identifier	HPOG-Impact site: PRS HPOG/PACE site: BIF
Employment in Healthcare	Ever employed in a healthcare job (binary) Note: This measure is not available for PACE programs.	Pre-intervention measure of outcome; Covariate	HPOG-Impact site: PRS
Hours	Number of hours worked for last week of employment (categorical)	Descriptor	HPOG-Impact site: PRS HPOG/PACE site: BIF
Hours Expectations	Expected working hours (categorical)	Descriptor	HPOG-Impact site: PRS HPOG/PACE site: SAQ
Employment Expectations	Expect to be working for pay in the next few months (binary)	Covariate	HPOG-Impact site: PRS HPOG/PACE site: SAQ
Participation Expectation	Expect to participate in HPOG full-time or part-time (categorical)	Descriptor; Subgroup Identifier	HPOG-Impact site: PRS HPOG/PACE site: SAQ

Notes: The measure capturing proportion of quarters employed during the four quarters prior to the quarter of random assignment comes from NDNH data.

Exhibit 2.5: Baseline Life Challenges Measures

Construct	Variable Description	Purpose(s)	Data Source(s)
Limited English proficiency	Limited English proficiency (binary)	Descriptor; Covariate; Subgroup Identifier	HPOG-Impact site: PRS HPOG/PACE site: BIF
Barriers to employment	Barriers to school/work (% at fairly or very often), including child care arrangements, transportation, health (series of binary variables)	Descriptor	HPOG-Impact site: PRS HPOG/PACE site: SAQ
Barriers to employment	Number of barriers that fairly often or very often interfere with school, work, job search or family responsibilities (integer measure, transformed to series of binary variables to identify subgroups)	Pre-intervention measure of outcome; Covariate; Subgroup Identifier	HPOG-Impact site: PRS HPOG/PACE site: SAQ

2.2.2 Outcome Measures

This section distinguishes among three types of outcomes: confirmatory, secondary and exploratory. Confirmatory outcomes are the main indicators of the extent to which the program is making progress towards its goals. Secondary outcomes are additional important outcomes identified in the HPOG logic model. Exploratory outcomes are of two types: (1) outcomes of interest that may be affected by the program but are not identified in the logic model and (2) alternative specifications of confirmatory and secondary outcomes.

Exhibit 2.6 lists the confirmatory and secondary outcomes to be included in the 15-Month Follow-Up Report. These outcomes reflect our preferred specifications of the outcomes pre-specified in the HPOG logic model and measure the following constructs:

- Educational Progress
- Earnings
- Employment
- Employment in Healthcare
- Employment Benefits
- Public Assistance Benefits

We consider our specifications of these outcomes to be a public commitment to planned analyses. This commitment guarantees that the choice of outcome and specification is made prior to the analysis of results and prevents intentional or unintentional data-mining or bias in selecting outcomes to analyze and present. Further details of operationalization for all outcomes, including specific data elements, appear in Appendix A.

Exhibit 2.6: Confirmatory and Secondary Outcomes at 15-Month Follow-Up

Domain	Variable Description	Outcome Designation	Data Source(s)
Educational Progress	Completion of training or ongoing enrollment in training	Confirmatory	15-month follow-up survey
Earnings	Wages received during the 5th quarter after quarter of random assignment[a]	Secondary	NDNH
Employment	Employed during the 5th quarter after quarter of random assignment[a]	Secondary	NDNH

Domain	Variable Description	Outcome Designation	Data Source(s)
Employment in Healthcare	Currently employed in a healthcare job or (if unemployed) worked for pay at some point after random assignment and most recent job was in healthcare	Secondary	15-month follow-up survey
Job Benefits	Current or most recent job offers health insurance	Secondary	15-month follow-up survey
Public Assistance Benefits	Individual receipt of cash public assistance (TANF) in the prior month	Secondary	15-month follow-up survey

Notes:
[a] We measure earnings and employment in the fifth quarter after random assignment. While participants are enrolled in training, they may forgo employment and earnings to focus on training. Measuring these outcomes in the fifth quarter after random assignment allows participants time to begin to realize any benefits of their training.

Exhibit 2.7 presents two confirmatory outcomes for the 36-Month Follow-Up report. The confirmatory hypotheses in this report address the overall impact of HPOG on these outcomes. Although the main purpose of this Analysis Plan is to specify the details of the 15-month follow-up analysis, we also identify the most important outcomes for later follow-up. Variables to be used for other analyses will be detailed in an analysis plan developed and released by the Career Pathways Intermediate Outcomes (CPIO) Study. Further details of operationalization for all outcomes for HPOG-Impact, including specific data elements, appear in Appendix A.

Exhibit 2.7: Confirmatory Outcomes at 36-Month Follow-Up

Domain	Variable Description	Outcome Designation	Data Source(s)
Educational Progress	Completion of training	Confirmatory	15-month follow-up survey 36-month follow-up survey
Earnings	Wages received during the 12th quarter after quarter containing random assignment	Confirmatory	NDNH

Next, Exhibit 2.8 lists the additional outcomes to be included in the 15-Month Follow-Up Report. These exploratory outcomes include alternative measures of some of the same domains measured by the confirmatory and secondary outcomes presented above, as well as measures in additional domains, as follows:

- Barriers to employment
- Economic Status
- Education and educational progress
- Employment
- Self-Efficacy & Motivation

Measurement of exploratory outcomes is intended to capture potential effects of the program that are less central to the program model and to inform new hypotheses. As such, we include the specification of these outcomes to share our thinking with the broader research community, but we do not view this list as set in stone. If our pre-specified analyses yield results that could benefit from additional exploration, we may add exploratory outcomes. Further details of operationalization for all outcomes, including specific data elements, appear in Appendix A.

Exhibit 2.8: Exploratory Outcomes at 15-Month Follow-Up

Domain	Variable Description	Outcome Designation	Data Source(s)
Barriers to Employment	Number of barriers that very often interfere with school, work, job search or family responsibilities	Exploratory	15-month follow-up survey
Earnings	Cumulative wages received during the five quarters after the quarter containing random assignment	Exploratory	NDNH
Economic Status	Personal income received from all sources	Exploratory	15-month follow-up survey
Economic Status	Household income received from all sources	Exploratory	15-month follow-up survey
Economic Status	Used loans in parents name to pay for school or living expenses	Exploratory	15-month follow-up survey
Economic Status	Used loans in either own name or parents name to pay for school or living expenses	Exploratory	15-month follow-up survey
Public Assistance Benefits	Personally received any government assistance in the prior month	Exploratory	15-month follow-up survey
Public Assistance Benefits	Number of major welfare programs (TANF, SNAP, Medicaid) from which the individual received benefits in the prior month	Exploratory	15-month follow-up survey
Public Assistance Benefits	Household received ANY government assistance in the prior month	Exploratory	15-month follow-up survey
Education	Completed a college degree (Associate's, Bachelor's or higher)	Exploratory	15-month follow-up survey
Education	Earned any college credits since random assignment	Exploratory	15-month follow-up survey
Education	Obtained a professional, state or industry certificate, license or credential since random assignment	Exploratory	15-month follow-up survey
Education	Completed a degree (AA, BA or higher) or obtained a credential (professional, state or industry certificate, license or credential) since random assignment	Exploratory	15-month follow-up survey
Educational Progress	Perception of progress towards long-range educational goals	Exploratory	15-month follow-up survey
Employment	Proportion of quarters employed during the five quarters after the quarter containing random assignment	Exploratory	NDNH
Employment	Time trend of employment, whether employed in each of the five quarters after the quarter containing random assignment	Exploratory	NDNH
Employment in Healthcare	Currently employed in a healthcare job	Exploratory	15-month follow-up survey
Self-Efficacy & Motivation	General Self-Efficacy Scale (GSE) based on Schwarzer & Jerusalem (1995)	Exploratory	15-month follow-up survey

2.3 Response Rates and Attrition

We expect administrative data coverage to be nearly complete, and our target response rate for survey data is 80 percent. That said, even well-designed randomized experiments may experience patterns of sample attrition that compromise the comparability of the treatment and control groups, potentially leading to biased estimates of the intervention's effectiveness (IES, 2014). We plan to conduct attrition analyses by computing both overall attrition (i.e., the rate of attrition for the entire sample) and differential attrition (i.e., the difference in the rates of attrition for the treatment and control groups).

Following the standards of the Clearinghouse for Labor Evaluation and Research (CLEAR) and the What Works Clearinghouse (WWC), these analyses will include all individuals who were randomly assigned.[3]

Because attrition includes individuals who are omitted from an analysis due to missing outcome data, each analytic sample will have a unique attrition rate. Attrition analyses will be performed for each analysis: the overall impact of HPOG, impact of HPOG on exogenous subgroups, and the impact of randomly assigned enhancements.

We plan to report in the main text overall attrition rates and differential attrition rates for each confirmatory hypothesis test. In an appendix, we will include more detailed tables with sample sizes by experimental group, the overall attrition rates, attrition rates by experimental group, and differential attrition rates associated with secondary hypothesis tests. Appendix B provides further details of our plan for calculating and reporting attrition.

2.4 Treatment of Missing Data

We anticipate encountering a variety of types of missing data. There will be individuals lost to survey follow-up (survey or unit nonresponse), individuals who refuse some questionnaire items or supply "don't know" responses (item nonresponse), and individuals with missing administrative data on HPOG participation. In addition, there will be a small number of individuals for whom administrative data from NDNH is missing.

This section describes our approach to missing NDNH, survey outcome data and individual baseline characteristics. Section 4.3 describes our approach to missing data on program characteristics and individual program participation measures.

2.4.1 Missing NDNH Outcome and Baseline Data

Data from the NDNH are used to construct outcome measures capturing earnings and employment, as well as baseline measures of the same constructs. In the NDNH data, we observe individual quarterly earnings from state Unemployment Insurance (UI) records, including data from some employers not included in the UI program (e.g., the federal government). Generally, individuals for whom we do not observe quarterly earnings in a particular quarter were not employed in that quarter. However, some of these individuals may have been employed and the observations missing due to issues matching administrative records.

Each quarter, the Office of Child Support Enforcement (OCSE), which maintains the NDNH data, submits HPOG sample members' social security numbers (SSN) and names to the Social Security Administration for verification before using these identifiers to match to the NDNH database. The SSN and name combinations not verified are eliminated from the match process.

Although NDNH output lists individuals with verification errors each quarter, these data do not allow us to perfectly distinguish individuals with missing earnings from the unemployed. The results of this

[3] Generally, the WWC treats all sample loss after random assignment as attrition. One key exception is that the WWC does not treat sample exclusions based on exogenous characteristics or characteristics measured prior to random assignment as attrition (IES, 2014). For analyses of the impact of HPOG on exogenous subgroups, we will include all individuals in the subgroup of interest who were randomly assigned as the base sample size.

verification are not consistent from one quarter to the next.[4] The match process for a later quarter updates earnings data from prior quarters. Because of this, we observe quarterly earnings for individuals with verification errors in the relevant quarters.

To distinguish between unemployment and missing data, we make the following assumptions:

- If an individual appears in the quarterly wage file for any quarter, we assume that quarters for which we do not observe earnings reflect periods of unemployment.
- If an individual never appears in the quarterly wage file and he or she does not appear on the list of verification errors, we assume he or she was unemployed for all quarters.
- If an individual never appears in the quarterly wage file and he or she appears on the list of verification errors in any quarter, we treat the NDNH derived earnings and employment data as missing for all quarters.

Preliminary analyses suggest that these assumptions are reasonable. We compared PRS measures of baseline characteristics for individuals for whom we treat data as missing to the characteristics of the two other groups listed above. The individuals for whom we treat data as missing appear more similar to the individuals for whom we observe quarterly wage data than to the individuals who were unemployed throughout the observation period. This suggests that data for individuals who never appear in the quarterly wage file and appear on the list of verification errors are better thought of as missing than as reflecting unemployment for all quarters. Applying these assumptions yields a missing rate of less than 5 percent.

We will use multiple imputations to address missing data from NDNH. Multiple imputations involves estimating an imputation model and using it to impute multiple possible values for each missing value. Compared with single imputation, multiple imputations allows for valid statistical inference in the impact analysis by better accounting for the uncertainty in the imputed values.

2.4.2 Missing Survey Data

For missing observations from follow-up survey data, we plan to use a combination of imputation and reweighting:

- *Imputation will be used to impute missing outcomes due to item nonresponse to the survey*. For sample members that completed part of the follow-up survey, baseline variables and completed follow-up survey items will be used to impute missing outcome values.
- *Reweighting will be used to address missing outcomes due to unit nonresponse to the survey.* For sample members that did not respond to the follow-up survey, baseline variables will be used to construct nonresponse adjusted weights.

To impute missing survey outcomes, we will use multiple imputations. Note that for our attrition calculations described in Section 2.3, missing outcomes will be counted as attrition even if that outcome is imputed.

[4] OCSE indicated that SSN verification status can change from "verified" to "non-verifiable" or vice versa for multiple reasons. Two examples are (1) when the name information submitted in a particular quarter is incomplete and (2) when individuals change their names due to marriage or divorce the SSA database and submitted information may cease to be aligned.

2.4.3 Missing Data Collected Via the PRS

Although some baseline measures are obtained from NDNH data, most baseline covariates were collected by the study through the PRS in HPOG Impact programs and via the BIF and SAQ in HPOG/PACE programs. We plan to use the same multiple imputation approach to adjust for item nonresponse in individual measures of baseline characteristics collected during intake for the study.[5] Unit nonresponse is not a concern for baseline measures: completion of these forms was required prior to random assignment.

Three proposed baseline covariates—past employment in healthcare, literacy, and numeracy— are not available for sample members in the HPOG/PACE programs.[6] Because these observations will be missing systematically, we will conduct an analysis to determine whether imputing these covariates for PACE sites or dropping the covariate entirely would be more appropriate.[7]

PRS data on program participation is administrative data that flags individuals who participate in specific training activities and supportive services. Individuals who are not flagged as participating may in fact not have participated, or the data on their participation may be missing. Because it is not possible to distinguish between the two scenarios, we will treat the data as non-missing and evidence that the individual did not participate.

2.4.4 Applying Nonresponse Analysis Weights

To help guard against imbalances caused by attrition, we plan to apply weights that adjust for survey nonresponse for analyses of outcomes collected from follow-up surveys.[8]

Note that applying nonresponse weights adjusts only for differences in observed characteristics between respondents and nonrespondents, separately for the treatment and control groups; it cannot adjust for differences in unobserved characteristics. To test whether the nonresponse analysis weights adequately control bias due to unit nonresponse, we propose to use outcome measures from NDNH data. Because the NDNH data have relative few missing observations, the main impact estimates for these outcomes are subject to relatively little nonresponse bias. We will test for any residual nonresponse bias by comparing the estimated impacts on NDNH earnings for the full sample to the estimated impacts on earnings for survey respondents, weighted with our nonresponse-adjusted weights. Small differences will suggest that the weighting adjustment is performing well.

If this test of the nonresponse analysis weights demonstrates that nonresponse biases impact estimates, even with the weighting adjustment, we will revisit model specification. We may be able to reduce nonresponse bias by improving the predicted probability of nonresponse by including a richer set of baseline variables from the PRS and outcome variables from administrative sources in the prediction model.

---

[5] Appendix C presents the rate of item-level nonresponse for baseline descriptors. These rates range from 0 to 20 percent.

[6] The missing observations for HPOG/PACE sample members are not reflected in the missing data rates presented in Appendix C as the denominator is restricted to the set of individuals for whom the variable is observable. Including the HPOG/PACE sample members in the denominator, the missing data rates for past employment in healthcare, literacy, and numeracy are 14, 34 and 37 percent respectively.

[7] In these analyses, we will explore the extent to which the covariates explain a significant portion of the variance in control group outcomes to determine if the covariates are worth retaining. We will also explore the extent to which we believe that the model predicting the missing values applies to the HPOG/PACE sites.

[8] We will follow Hsueh et al. (2012) and Izrael, Battaglia, and Frankel (2009) in our construction of nonresponse weights.

3. Description of the Study Sample

This section describes the study sample at intake on a wide range of measures. These tables include all individuals who were randomly assigned and provided consent for data collection.[9]

Random assignment ensures that there are no systematic differences in baseline characteristics between the groups, though differences can emerge due to chance. These exhibits identify baseline characteristics for which the differences are statistically significant at the 5-percent level. Significant differences do not indicate a failure of randomization; we would expect to find significant difference for 5 percent of the baseline variables due to random chance.

The sample is overwhelmingly female (89 percent) and the majority had never married (61 percent) or had children (61 percent) at baseline. One-third of the sample is Non-Hispanic White, another third is Non-Hispanic African American, and one-fourth is Hispanic. About 18 percent of the sample reported being born outside the U.S. In terms of education, 43 percent of the sample had completed high school or an equivalent at baseline and more than half had attended college or post-secondary training. More than one-fifth had an occupational skills certification or license in hand. Despite these education levels, most sample members had low income at baseline: 43 percent had less than $10,000 in household income, and 22 percent had no individual income at all. At baseline, 14 percent were receiving welfare assistance and 57 percent were receiving government food assistance (SNAP or WIC). In terms of life challenges, 86 percent of the sample demonstrated literacy at the eighth grade level or above and 76 percent numeracy at the eighth grade level or above. Relatively few report meaningful barriers to their participation in school or work: about 14 percent report either child care or transportation as barriers, and about 10 percent report that a health condition poses a challenge.

As the exhibits show, the differences between the groups in these characteristics are small in magnitude. Of 73 tests conducted at the 0.05 level of significance, the null-hypothesis of no difference between the treatment and comparison groups was rejected 17 times. We expect to see three or four statistically significant differences given 73 independent tests at the 0.05 level of significance when the null-hypothesis was true. We have multiple possible explanations for the larger number of statistically significant differences, all of which likely play some role in the number of significant differences.

First, treatment group baseline data could have been updated after random assignment more frequently than control group data. Only a very small number of variables were required to be entered in order to carry out the random assignment process. Baseline data were reported via the PRS and could be updated by program staff throughout data collection, although programs were explicitly instructed not to do this. Investigations have found a very small number of cases in which baseline data were updated. However, we do not believe that these few cases could introduce meaningful differences between treatment and control groups.

Second, systematic differences in the rates of missing data may produce statistically significant differences in observed measures in cases where the groups are actually balanced. Missing data rates are presented in Appendix C. Of the 17 apparently statistically significant differences, 11 are calculated based on measures with statistically significantly different missing data rates between treatment and control

[9] A total of 236 individuals were randomly assigned and did not provide consent or withdrew consent. These individuals are not included in these descriptive statistics, but will be included in the baseline sample sizes for attrition calculations.

groups.[10] This suggests that differences in missingness may be part of the explanation for the apparent differences between the groups.

Alternately, when constructs are highly correlated, the number of statistically significant differences could be larger in number than would be expected by chance alone. Nine of the 17 measures that show statistically significant differences measure educational background and expectations, and the remaining eight measure demographic characteristics. For example, the between-group differences in racial and ethnic composition represent five of the observed statistically significant differences; but these are highly correlated, small in magnitude, and relate to one construct, reducing our concern about the number of them.

Finally, it is possible, though unlikely, for this number of null-hypotheses to be falsely rejected due to randomly occurring differences between the treatment and control group samples (but no underlying differences between the populations from which they were selected).

Although the number of significant findings exceeds expectation, the very small magnitudes of the differences are reassuring. For example, though the difference in the racial and ethnic distribution of the treatment and control groups at baseline is statistically significant in five of the seven categories, it does not appear to be meaningful in magnitude. In the treatment group, 36.4 percent of individuals are White/Caucasian and 34.7 percent are Black/African American; whereas, within the control group, the corresponding numbers are 34.4 and 32.4, respectively. These represent about only a two percentage point difference in the prevalence of these racial identities between the two groups. As such, we do not believe that these differences are substantively important; nevertheless, we include baseline measures in the model to account for these differences as well as other random baseline differences between the two groups.

Exhibit 3.1 presents the number of grantees, programs, divisions and individuals included in the sample. The sample sizes and rate of missing data associated with each characteristic in Exhibits 3.2 to 3.6 are presented in Appendix C.

Exhibit 3.1: Sample Size at Baseline

Characteristic	Entire Sample	Treatment Group	Control Group
Sample Size			
Grantees (N)	23	23	23
Programs (N)	42	42	42
Administrative Divisions (N)	87	87	87
Individuals (N)	13,575	8,601	4,974

[10] Missing data rates differ at the 0.05 level of significance for Race, Born outside the U.S., Education level expected to complete, and Numeracy at the 8th grade level or higher. Rates of missing data were not significantly different at the 0.05 level for Age, Post-secondary degree/certificate completion, Previous preparation classes in how to succeed in work and Expect to participate in HPOG full vs. part time.

Exhibit 3.2: Demographic Characteristics of Sample at Baseline

Characteristic	Entire Sample	Treatment Group	Control Group	Significant Difference
Sex				
Male (%)	11.2	11.0	11.6	
Female	88.8	89.0	88.4	
Marital Status				
Married (%)	16.4	16.1	16.8	
Separated or divorced (%)	21.8	21.3	22.7	
Widowed (%)	1.2	1.2	1.3	
Never Married (%)	60.6	61.3	59.2	**
Dependent Children				
None (%)	38.6	38.7	38.5	
One or two (%)	45.5	45.2	45.9	
Three or more (%)	15.9	16.1	15.6	
Race and Ethnicity[a]				
Hispanic/Latino of any race (%)	23.6	22.5	25.4	**
White/Caucasian (%)	35.7	36.4	34.4	**
Black/African-American (%)	33.9	34.7	32.4	**
Asian, Native Hawaiian, Pacific Islander (%)	3.3	3.0	3.8	**
American Indian or Native Alaskan (%)	0.8	0.7	1.1	**
Two or more races (%)	2.7	2.6	2.9	
Age				
Less than 20 years (%)	7.6	7.4	7.9	
20-24 years (%)	24.5	25.5	22.8	**
25-29 years (%)	19.1	19.0	19.1	
30-34 years (%)	14.6	14.4	14.9	
35-39 years (%)	10.3	10.0	10.7	
40-49 years (%)	14.4	14.3	14.6	
50+ years (%)	9.6	9.4	9.9	
Born Outside the U.S. (%)	17.8	17.2	18.7	**

Notes:
** Indicates difference between treatment and control group: $p<0.05$, two-tailed t-test of difference in means
[a] The categories White/Caucasian, Black/African-American, Asian, Native Hawaiian, Pacific Islander, American Indian or Native Alaskan, and Two or More Races only include individuals who do not identify as Hispanic/Latino.

Exhibit 3.3: Educational Background of Sample at Baseline

Characteristic	Entire Sample	Treatment Group	Control Group	Significant Difference
Educational Attainment				
Less than 12th grade (%)	5.4	5.2	5.8	
High school equivalency/GED (%)	13.3	13.2	13.4	
High school graduate (%)	30.3	30.9	29.3	
1-3 years of college/tech school (%)	42.1	41.9	42.4	
4 years or more of college (%)	8.9	8.8	9.1	
Credential/Degree completion				
Post-secondary degree/certificate (%)	25.7	24.5	27.6	**
Occupational skills license or certification (%)	22.2	22.2	22.2	
Literacy and Numeracy				
Literacy at the 8th Grade Level or Higher[a] (%)	85.8	85.7	85.8	
Numeracy at the 8th Grade Level or Higher[a] (%)	75.9	75.2	77.5	**
Previous Preparation Classes				
Adult basic education (%)	17.6	17.8	17.3	
English as a second language (%)	8.0	7.8	8.4	
Classes in how to succeed in school (%)	15.5	15.7	15.2	
Classes in how to succeed in work (%)	16.5	16.0	17.3	**
Vocational, technical or trade school (%)	31.1	31.0	31.3	
Level Expected to Complete				
No additional (%)	1.7	1.6	1.8	
GED (%)	4.7	4.9	4.3	
High school diploma (%)	9.8	10.2	9.1	**
Alternative non-academic credential (%)	12.4	13.1	11.4	**
Associate's degree (%)	24.5	23.9	25.6	**
Bachelor's degree (%)	30.0	29.3	31.2	**
Graduate degree (%)	16.8	17.0	16.6	

Notes:
** Indicates difference between treatment and control group: $p<0.05$, two-tailed t-test of difference in means
[a] Literacy and Numeracy were not available for HPOG/PACE participants.

Exhibit 3.4: Economic Status of Sample at Baseline

Characteristic	Entire Sample	Treatment Group	Control Group	Significant Difference
Household Income[a]				
$9,999 or less (%)	43.1	43.3	42.7	
$10,000 to $14,999 (%)	16.2	16.2	16.3	
$15,000 to $19,999 (%)	12.2	12.4	11.7	
$20,000 to $29,999 (%)	16.2	16.2	16.1	
$30,000 to $39,999 (%)	6.8	6.6	7.1	
$40,000 or More (%)	5.6	5.4	6.0	
Individual Income[b]				
$0 (%)	22.2	22.5	21.5	
$1 to $9,999 (%)	41.6	41.0	42.8	
$10,000 to $14,999 (%)	14.1	14.3	13.7	
$15,000 to $19,999 (%)	9.5	9.5	9.5	
$20,000 to $29,999 (%)	9.8	9.9	9.5	
$30,000 or More (%)	2.8	2.7	3.0	
Public Assistance Use				
WIC/SNAP (% receiving)	57.4	56.9	58.3	
Welfare (% receiving)[c]	13.6	13.8	13.3	

Notes:
** Indicates difference between treatment and control group: p<0.05, two-tailed t-test of difference in means
[a] HPOG/PACE participants reported family income, whereas HPOG/Impact participants reported household income.
[b] HPOG/PACE participants did not report individual income.
[c] HPOG/PACE participants reported whether they received any form of welfare, whereas HPOG/Impact participants reported only TANF.

Exhibit 3.5: Employment and Expected Time Use of Sample at Intake

Characteristic	Entire Sample	Treatment Group	Control Group	Significant Difference
Current Working Hours				
Not working (%)	57.8	57.4	58.4	
<20 hours per week (%)	11.6	11.5	11.6	
20-34 hours per week (%)	18.5	18.5	18.5	
35+ hours per week (%)	12.2	12.6	11.5	
Expected Working Hours				
Not working (%)	58.9	58.8	59.2	
<20 hours per week (%)	6.7	6.7	6.9	
20-34 hours per week (%)	17.7	18.0	17.1	
35+ hours per week (%)	16.7	16.6	16.8	
Expect to Participate in HPOG				
Full-time (%)	74.4	73.8	75.5	**
Part-time (%)	25.6	26.2	24.5	**

Notes:
** Indicates difference between treatment and control group: p<0.05, two-tailed t-test of difference in means

Exhibit 3.6: Life Challenges of Sample at Baseline

Characteristic	Entire Sample	Treatment Group	Control Group	Significant Difference
Limited English Proficiency (%)[a]	1.0	1.1	0.9	
Barriers to school/work (% at fairly or very often)				
Child care arrangements (%)	14.3	14.4	14.1	
Transportation (%)	14.8	14.9	14.6	
An illness or health condition (%)	10.0	10.2	9.8	
Alcohol or drug use (%)	0.6	0.5	0.8	

Notes:
** Indicates difference between treatment and control group: p<0.05, two-tailed t-test of difference in means
[a] Limited English Proficiency is defined as speaking English "not well" or "not at all" for HPOG/PACE participants. It is defined as directly for HPOG only participants.

4. Measures of Program Characteristics and Participation

In this section, we present measures related to program characteristics. These measures are used in the analysis of the role of program components described in Section 6. In addition, we present measures of program participation at the individual level. These measures are used in the analysis of the role of participation in particular program components in determining the effect of the program, as described in Section 7.

4.1 Program Characteristics and Context

We distinguish among the following types of program characteristics and context measures:

- *Program components* describe the services available to HPOG participants. These data are measured at the program level based on the National Implementation Evaluation (NIE) Grantee Survey.[11]
- *Implementation features* describe how services are delivered to HPOG participants. These data are measured at the division level based on responses to the NIE Management and Staff Survey.
- *Local context variables* measure the larger economic environment within which a particular HPOG grantee is situated. These indicators are constructed from data publically available from the Bureau of Labor Statistics and the Census Bureau and are measured at the program level.[12]
- *Participant composition measures* are division-level aggregations of individual-level baseline characteristics. The candidate measures are covariates listed in Exhibits 2.1–2.7.

We select these specific program characteristics and context measures because we expect them to be related to the effectiveness of the program. Below, we discuss our understanding of and hypotheses regarding the relationship between these measures and the impact of the program.

4.1.1 Program Components

A description of each program component measure appears in Exhibit 4.1. These program components include both the three enhancements that ten of the grantees are testing through three-arm random assignment; they also include other enhancements that vary naturally across grantees. In the text we provide the motivation for each measure as well as the hypothesized link between measures and impact.

Presence of Career Pathways Principles

We will to measure the extent to which available offerings and program content are based on the principles of the Career Pathways framework. Career Pathways (CP) is a framework of services and basic academic and vocational training strategies intended to assist low-income individuals and other nontraditional post-secondary students in completing increasingly complex courses articulated by specific sectoral skill requirements. This measure was developed to align with the Career Pathways framework articulated by Fein (2012). In in the process of creating the measure, we investigated the extent to which program implementation reflected the principles of the Career Pathways framework through qualitative

[11] The NIE aims to describe and assess the implementation, systems change, and outcomes related to the 27 non-tribal HPOG grantees.

[12] Many observations for programs within the same grantee have the same value of these measures, because these programs experience the same economic environment. Appendix A provides more detail on how HPOG programs are mapped to Metropolitan Statistical Areas to obtain local context measures for programs and grantees from Census and BLS data sources.

coding of information collected via implementation site visits. This qualitative coding exercise identified the same constructs that we ended up including in the measure, drawing on quantitative survey data. The qualitative data analysis provides additional validation of this specific construct as appropriate for this project's use.

Although CP has become increasingly popular as a strategy for increasing the employment and wages of low-income individuals and is believed to have improved the outcomes of participants in selected programs, its impacts have not yet been tested in a rigorous evaluation.[13] This "presence of CP principles" construct will be used to assess the degree to which greater alignment with the CP framework affects HPOG program impacts.

Exhibit 4.1: Program Components

Domain	Variable Description
Presence of Career Pathways Principles	Extent to which available offerings and program content is based on principles of the career pathways framework
Case Management	Average caseload for FTE (estimated full time equivalent) case managers
Case Management	Number of services that case managers and counselors deliver that meet the needs of participants
Comprehensive Services	Access to social and other services: social and other services delivered that meet participants' needs
Comprehensive Services	Access to and delivery of tuition and other financial services: tuition coverage plus financial services offered that meet participant needs
Comprehensive Services	Access to childcare and transportation: accessibility via public transportation plus childcare and transportation services offered that meet participant needs
Comprehensive Services	Location of services: number of services co-located with the training site
Employment Supports	Number of employment supports that are offered that meet participants' needs
Behavioral Incentives	Non-cash incentives: whether the program provides offer non-cash incentives to participants for achieving program milestone
Peer Support	Offer of facilitated peer support
Emergency Assistance	Access to emergency funds to meet needs stemming from imminent eviction from housing, utility shutoff, vehicle repair needs, etc.

Case Management

We will measure average caseload and the number of services that case managers and counselors deliver. Case managers and counselors are the frontline staff most directly in contact with participants and most directly responsible for meeting participant needs and fostering program completion and employment success. We hypothesize that lower caseloads will lead to greater access to case managers and their services for participants. Indeed, prior research has found that lower caseloads are correlated with larger impacts on employment in an evaluation of a national welfare reform program (Bloom, Hill & Riccio, 2003).

HPOG programs have made varying design decisions around the choice of "generic" vs. "specialized" case management. Some programs use more generic case managers (under various titles) who may provide basic case management as well as a variety of personal, financial, academic, and employment counseling. Other programs use specialist case managers or counselors for each type of support service. The "generic" approach may be more convenient for participants and may lead to a deeper and more personal relationship between case manager and student. Bloom, Hill, and Riccio (2003) found a stronger personal relationship between case manager and client to be associated with larger impacts on

[13] For a review of the evaluation literature on Career Pathways, see Werner et al. (2013).

employment. Similarly, research on the effects of an enhanced student services program at two community colleges found that developing a personal relationship with non-academic staff is associated with students' higher educational success (Scrivener & Weiss, 2009). On the other hand, specialists may be more skilled and effective in their specific roles. There are merits for each approach, with no strong empirical evidence for either approach. We hypothesize that programs where case managers and counselors provide a greater number of services are less likely to provide specialized case management. Including a measure of the number of services that case managers and counselors deliver holds the potential to develop a recommendation to the field.

Comprehensive Services

In the comprehensive services domain, we measure access to social and other services. By intent, HPOG serves a low-income population that faces many barriers to sustained enrollment, educational attainment, and career advancement. To complete training and compete successfully in the healthcare labor market, HPOG participants require a variety of personal, academic, social, financial supports, and other services.

Academic supports encompass the range of services that focus on academic needs, and may include individual tutoring, group sessions on specific academic or vocational topics, study groups, and self-paced computerized instruction (Jobs for the Future, 2010; Stephens, 2009). Academic supports may also include training that supplements vocational training, such as college-readiness training. College-readiness training is intended to provide nontraditional postsecondary students with an understanding of expectations and responsibilities of students, and strategies for navigating and completing postsecondary education (Karp, 2011). This may be especially important to the educational success of the HPOG population, of who almost half have no previous experience with postsecondary education.

Financial assistance may facilitate academic participation and completion for HPOG participants. Financial constraints is one of the most commonly cited barriers to low-income students' entering and completing post-secondary education, according to the Education Longitudinal Study of 2002 (ELS, 2002). All programs provide some financial assistance, ranging from full tuition waivers to partial tuition assistance, as well as assistance for academic supplies, exam fees, uniforms, etc. Relieving more of the financial burdens that the low-income participants experience may lead to greater success. This measure of access to and delivery of tuition and other financial services measure allows the study to test whether perceived adequacy of and more direct access to financial services is related to the impact of HPOG.

Childcare and transportation assistance may also be critical supports for HPOG's low-income population, of whom over 60 percent have dependent children. The access to childcare and transportation measure allows the study to test whether the perceived adequacy of and more direct access to these services relate to the impact of HPOG.

We also measure the perceived adequacy of, and more direct access to other support services such as non-SNAP food assistance, family preservation services, legal services, etc., tests the widely held belief about their relationship to program success (Estrada, 2010; Hinckley & Hull, 2009; Jobs for the Future, 2010).

Services included in the location of services measure are those typically provided through in-person meetings with a case manager or counselor. The study has learned from our site visits that HPOG management and staff generally believe that students will use those services more readily if they are conveniently located. This measure allows a test of the hypothesis that physical ease of access to personal, academic, financial and employment services is associated with larger impacts on program completion and employment.

Employment Supports

Prior research and field wisdom has shown that various approaches to employment assistance—such as job search training, job-readiness training, job fairs—can be help low-income workers find employment. The CP literature states that employers should provide students internships or other work-based learning opportunities (Estrada, 2010; Pusser & Levin, 2009; Soares, 2010). These opportunities are expected to improve participants' career awareness and knowledge, facilitate connections with local employers, and build participants' resumes while they are in training (Fein, 2012). With the employment supports measure, the study can determine if providing more types of assistance is associated with larger impacts on employment.

Non-cash Incentives

The post-secondary education community has used both results- and behavior-based incentives to increase school attendance, completion and performance. Although there is no reliable evidence on the impact of non-cash incentives, CP literature suggests that awarding students with financial incentives is a promising strategy that both celebrates students' accomplishments (Endel, Anderson & Kelly, 2011) and addresses students' financial needs (Kazid & Liebowitz, 2003). The Opening Doors study found that cash incentives increase the attainment of college credits, school retention and motivation (Richburg-Hayes et al., 2009). A study of a the New Hope program, an incentive program that offers financial and non-financial work supports to working poor families in Milwaukee, WI found that the program reduced by half the number of families who were never employed during the study period. Program participants' earnings were also 13 percent higher than nonparticipants' earnings (Bos et al., 1999). In the HPOG context, non-cash incentives allow students/trainees to earn points for achieving specific program milestones and convert those points into tangible rewards, such as vouchers for use at the college bookstore, work-related equipment (such as scrubs or a stethoscope), or gift cards to support meeting basic needs (such as for transportation/gas or food). This measure allows us to test the association of non-cash incentives with impacts on a broader range of outcome measures.[14]

Peer Support

Programs that offer a peer support structure provide opportunities for trainees/students to create personal relationships that may increase their accountability and commitment to program retention and completion. For example, Karp (2011) reports more-favorable outcomes among students that established meaningful social relationships; and Grant-Vallone et al. (2004) find that better-adjusted students are more committed to their educational goals. Non-traditional students tend to have lower levels of social integration (Tinto, 1993), something peer support may help remedy. This evidence is suggestive and highlights that a more rigorous test is justified. Anecdotal evidence from HPOG programs and other training programs for low-income populations suggest that strong peer connections can foster greater program attachment and group identification. This measure tests the field wisdom that peer support groups can lead to better rates of course and program completion.[15]

Emergency Assistance

Although HPOG programs include a wide variety of specific financial supports, not all programs cover all participants' unforeseen needs that may interfere with course attendance, such as car repairs or housing crises. Some programs provide assistance for a wider range of potential needs than others. In a 2008 study

[14] More information on non-cash incentives is available in Section 4.3.1 of the Evaluation Design Report on pages 43-44.

[15] More information on peer support is available in Section 4.3.1 of the Evaluation Design Report on page 42.

of emergency financial aid programs assisting community college students who are in danger of dropping out due to sudden financial crises, researchers reported that college administrators as well as students claimed that the emergency financial assistance helped students stay enrolled in college (Geckeler, 2008). This measure will permit testing the proposition that provision for a broader range of emergency needs is associated with larger impacts on program retention and completion.[16]

4.1.2 Implementation Features

Implementation features measures capture the way in which program services are delivered. These measures, which describe management and staff characteristics and perceptions, are presented in Exhibit 4.2 below.

Management/Staff Focus

HPOG programs must balance participants' more immediate need for employment with the goal of establishing the groundwork for longer-term career advancement. Prior research has demonstrated that an emphasis on quick job entry is associated with an increase in the effectiveness of training programs (Bloom, Hill & Riccio, 2003). However, the CP framework also emphasizes education "organized as a series of manageable steps leading to successively higher credentials and employment opportunities in growing occupations" (Fein, 2012). This measure allows the study to determine whether staff and management beliefs about the primary goal of HPOG (employment or education) are associated with different impacts on employment and earnings.

Staff Experience

Successful HPOG implementation and performance relies in large part on the skill and efficacy of program management and staff. The staff experience measure allows the study to test the hypothesis that case management divisions with more experienced HPOG staffs are associated with larger participant impacts.

Staff Discretion/Autonomy

Research on front-line human services workers has identified a variety of bureaucratic strategies used to implement human services programs (Brodkin, 1997; Evans, 2010). One dimension is the degree of autonomy afforded front-line workers in their interactions with program participants, ranging from closely monitored and supervised rule-dominated regimes to those that allow front-line workers greater discretion in those interactions. The measure of staff discretion/autonomy allows the study to determine whether greater perceived worker discretion at the division level leads to greater or smaller impact sizes.

Exhibit 4.2: Implementation Features

Domain	Variable Description
Management/Staff Focus	Extent to which program is employment or education focused
Staff Experience	Percentage of management/staff at the division level with at least five years of experience
Staff Discretion/Autonomy	Staff perception of autonomy, including authority to carry out responsibility, ability to try different techniques, trust in staff professional judgment and not too many rules.

[16] More information on emergency assistance is available in Section 4.3.1 of the Evaluation Design Report on page 42-43.

4.1.3 Local Context Variables

Exhibit 4.3 presents the local context variables. The local labor market conditions will affect HPOG trainees and control group members' ability to find employment and the amount they earn when employed (Bloom, Hill & Riccio, 2003). The effectiveness of a training program at increasing employment is closely linked to the local demand for the jobs for which participants are trained. These measures will capture local variation in the labor market conditions that may explain the relative effectiveness of HPOG programs in different locations.

Exhibit 4.3: Local Context Variables

Domain	Variable Description	Data Source(s)[a]
Population	Total Population	American Community Survey 5-Year Estimates (B01003)
Education	Percent of adult population age 25 and over with some college; associate's degree; bachelor's degree; or higher	American Community Survey 5-Year Estimates (B15003)
Education	Percent of adult population age 25 and over enrolled in school	American Community Survey 5-Year Estimates (B14003)
Receipt of Public Assistance	Percent of households that receive cash public assistance income, Food Stamps/SNAP, or both in the past 12 months	American Community Survey 5-Year Estimates (B19058)
Labor Market Conditions	Unemployment rate (across all occupations)	BLS Unemployment Statistics for States and Metropolitan Statistical Areas (MSAs)
Healthcare Specific Labor Market Conditions	Percent of jobs that are in healthcare sector	Bureau of Labor Statistics Occupational and Employment Statistics (OES) for MSAs
Healthcare Specific Labor Market Conditions	Median wage of healthcare support occupations	Bureau of Labor Statistics (OES) for MSAs

Note:
[a] We will use the most recent version of the data set available as of August 2016 for the Final Report to align the timeframe of these measures with the reference period for the outcome measures.

4.2 Individual-Level Program Participation Measures

Exhibit 4.4 lists measures capturing individual-level program participation that will be analyzed as mediators of impact (see Section 7). The analysis of program components may be thought of as investigating the impact of offering a service, while the analysis of program participation investigates the impact of taking up the offer. To allow us to follow the story of service-driven impact from offer to take-up, the specific program components were selected to align with the analysis of randomly assigned enhancements (emergency assistance, non-cash incentives, and peer support) and the analysis of program components that vary naturally (personal counseling and academic assistance).

Exhibit 4.4: Individual-Level Measures of Program Participation at 15-Month Follow-Up

Domain	Variable Description	Data Source(s)
Education	Obtained a professional, state or industry certificate, license or credential since random assignment	15-month follow-up survey
Education	Completed a degree (AA, BA or higher) or obtained a credential (professional, state or industry certificate, license or credential) since random assignment	15-month follow-up survey
Emergency Assistance	Received emergency assistance	PRS 15-month follow-up survey
Non-Cash Incentive	Received a non-cash incentive	PRS 15-month follow-up survey

Domain	Variable Description	Data Source(s)
Peer Support	Participated in peer-support	PRS 15-month follow-up survey
Personal Counseling	Used personal counseling services	PRS 15-month follow-up survey
Academic Assistance	Used academic assistance services (tutoring, etc.)	PRS 15-month follow-up survey

4.3 Treatment of Missing Aggregate-Level Data

As described in Section 2.4, we anticipate encountering a variety of types of missing data relevant to program measures. Here we address the aggregate level data and our plans for handling missing data therein. Specifically, we expect very low rates of item nonresponse in program component and implementation features data constructed at the program- and division-levels. The NIE conducted extensive follow-up to ensure low rates of missing data in the surveys from which we will construct these measures. Therefore, we will use a simpler approach for these data—case deletion—in instances where we cannot construct complete data.

5. Impact Analysis

This section provides an overview of the Research Questions, plan for hypothesis testing, and other considerations for the HPOG Impact Study's impact analyses. Specifically, Section 5.1 describes the Research Questions. Section 5.2 describes the HPOG Impact Study's approach to hypothesis testing. Section 5.3 explains how hypothesis test results will be presented in project reports. Section 5.4 provides the definitions of the arithmetic terms included in the analytic models presented in Sections 5 through 7. Section 5.5 then presents the analysis plan for addressing Research Question 1: What impacts do the HPOG programs as a group have on outcomes for participants and their families? Section 5.6 describes several potential approaches to conducting sensitivity analyses. Finally, Section 5.7 presents the analysis plan for answering Research Question 2: To what extent do impacts vary across selected subpopulations? The analysis approaches for Research Questions 3 and 4 are addressed in Sections 6 and 7, respectively.

5.1 Research Questions

The study's first Research Question concerns the HPOG program's overall effectiveness:

> *Research Question 1*: What impacts do the standard HPOG programs as a group have on outcomes for participants and their families?

The second Research Question concerns the possibility that impacts vary by segment of the target population:

> *Research Question 2*: To what extent do impacts vary across selected subpopulations? The analysis plan for answering Research Question 2 is described in detail in Section 5.5.

The third and fourth Research Questions relate to how variation in program design drives variation in impacts and can help us understand what about these programs makes them work:

> *Research Question 3*: Which locally adopted program components influence average impacts?
>
> *Research Question 4*: To what extent does participation in a particular HPOG component (or components) change the impact on individual trainees?

These last two questions differ in that one focuses on the program's design and the other focuses on the individual's experience in the program.

5.2 Approach to Hypothesis Testing

In addressing these questions we classify impact hypotheses as confirmatory, secondary, and exploratory. Confirmatory hypotheses are those that can inform the extent to which the program is making progress toward its goals. Secondary hypotheses are about additional relationships that are pre-specified in the HPOG logic model. Exploratory hypotheses are about program effects that may help improve our understanding of the findings from the confirmatory and secondary analysis.

Conducting tests of statistical significance for too many impact findings creates what is known as the "multiple comparisons problem." The likelihood of finding one or more statistically significant impacts purely by chance when many tests are conducted can be quite high—much higher than the 5 percent chance that an individual test would suggest. For example, if we were to conduct impact analyses on five outcomes that were unaffected by an intervention, the probability of finding one or more statistically significant impact estimate by chance when the odds of having falsely significant results are set at 5 percent for individual tests is 23 percent (assuming the tests are independent). For ten tests, the

probability rises to 40 percent. This risk is viewed as unacceptable for the main findings of hypothesis testing in social program impact evaluations. In response, we propose to adjust for multiple comparisons for our most important analyses.

Since multiple comparison adjustment is relatively recent in social policy evaluation, there are no universally accepted standards about when adjustment is necessary. Some experts in the field have adopted the standard that adjustment should occur whenever there is more than one confirmatory hypothesis, and others suggest that adjustments should be domain specific. The rationale for making domain-specific adjustments, and not making adjustments across domains, is that hypotheses do not span multiple domains. The strength of protection against Type I errors, then, may depend on how many domains are allowed to be defined.

To keep the risk of false positive findings close to the stated *p*-values, avoid the need for adjustment and attendant loss of power, and simplify interpretation, together the evaluation team and ACF decided that HPOG-Impact should follow the practice of limiting confirmatory tests. In response, we plan to conduct one confirmatory hypothesis test at the 15-month follow-up to see if on average, HPOG has a positive effect on educational progress. At 36-months, we plan one confirmatory hypothesis in each of two domains—educational progress and earnings—to gauge the program's success in improving individuals' circumstances. We will not adjust for multiple comparisons across the two confirmatory tests at the 36-month follow-up, with the justification that educational progress and earnings are different outcome domains, nor will we adjust for the multiple tests to be examined at two points in time. We will confine the confirmatory analysis to examination of the standard program and treat the experimental enhancement analyses, subgroup analyses, and mediator analyses of the same outcomes as secondary or exploratory.

Combined, the type of analysis and the type of outcome together determine whether a hypothesis test is classified as confirmatory, secondary or exploratory. Along these lines, we characterize *outcomes* as confirmatory, secondary and exploratory. These designations align with the level of the *hypothesis* test regarding the overall impact of HPOG on the outcome. Confirmatory hypotheses, which are designed to inform the extent to which the HPOG program is making progress toward its goals, address the impact of HPOG on the following confirmatory outcomes:

- Educational progress as measured by completion of our ongoing enrollment in training at 15-month follow-up,
- Educational progress as measured by completion of training at 36-month follow-up, and
- Earnings as measured by wages received by the 12th quarter after the quarter containing random assignment at 36-month follow-up.

Secondary outcomes are additional important outcomes identified in the HPOG logic model. At the 15-month follow-up, tests of the overall impact of HPOG on the following secondary outcomes are secondary hypotheses:

- Earnings measure constructed from NDNH data,
- Employment measure constructed from NDNH data,
- Employment in healthcare,
- Employment benefits, and
- Public Assistance benefits as measured by receipt of TANF in the prior month.

Confirmatory and Secondary outcomes are operationalized in Appendix A, Exhibits A.6 and A.7.

Exploratory outcomes are of two types: (1) outcomes of interest that may be affected by the program but are not identified in the logic model and (2) alternative specifications of confirmatory and secondary outcomes. These alternative specifications provide additional context for and aid in the interpretation of the main findings for the impact of HPOG on exploratory and secondary outcomes. All analyses of exploratory outcomes inform exploratory hypotheses. These outcomes are operationalized in Appendix A, Exhibit A.8.

We consider both the importance of the outcome and the importance of the analysis when determining if the level of evidence of a given hypothesis test is classified as confirmatory, secondary, or exploratory. Therefore, although outcome designations—confirmatory, secondary and exploratory—align with the level of the hypothesis test for analyses of the overall impact of HPOG, these designations do not align for the experimental enhancement analyses, subgroup analyses, and mediator analyses of the same outcomes. Exhibit 5.1 describes how each combination of type of analysis and outcome aligns with the level of evidence for hypothesis testing.

We do identify multiple secondary and exploratory hypotheses for outcomes in the same domain. As is standard practice,[17] we do not perform multiple comparisons corrections for these hypothesis tests. Therefore, their findings must be presented with appropriate caveats described in the next section.

The team limits the number of secondary hypothesis tests—those stated in this document—thereby containing the risk of false positive findings. However, the limitations associated with tests of secondary hypotheses will be less strict in that the team will allow a careful, but more generous, pre-specification of outcomes for secondary testing. As shown in the third column of Exhibit 5.1, secondary hypothesis testing (as distinct from secondary outcome designation in the second column) will consist of the following:

- Tests of overall program impacts on secondary outcomes;
- Tests of the overall HPOG impact on confirmatory outcomes for key subpopulations; and
- Tests of the impact of randomly assigned HPOG enhancements on confirmatory outcomes.

Exhibit 5.1 also indicates planned exploratory hypothesis testing in several rows. These consist of the following tests: tests of overall program effects on exploratory outcomes; most hypotheses concerning the overall HPOG impact on subpopulations; most hypotheses concerning mediation of program effects via program components; and all hypotheses related to variation in impacts due to various program participation patterns of individuals.

For all impact analyses, we will report three thresholds for statistical significance, each with a distinct meaning as concerns the strength of evidence: an alpha level of 0.10 (a 10-percent chance of concluding an impact has occurred when none has) will be used for "suggestive" evidence, 0.05 for "moderate" evidence, and 0.01 for "strong" evidence. As shown in the final column of Exhibit 5.1, confirmatory and secondary analyses of HPOG's overall impact will use one-tailed tests and all other analyses will use two-tailed tests. The reason for this is as follows: the confirmatory and secondary analyses have their foundations in the program's logic model, which implies a directional hypothesis and therefore justifies a one-tailed test. Other analyses, such as those concerning the relative effectiveness of the program on

[17] For example, Schochet (2008a) recommends that non-confirmatory hypotheses need not be subject to multiple comparisons corrections, provided the appropriate caveats to interpretation are provided.

various subgroups, do not have a clear directional hypothesis, and so it is more appropriate to use a two-tailed test. We discuss next the implications of this scheme for reporting results.

Exhibit 5.1: Outcomes and Hypothesis Tests for 15-Month Analysis

RQ	Estimate of Interest	Outcome Designation (Confirmatory/Secondary /Exploratory)	Level of evidence of hypothesis test (Confirmatory/Secondary/ Exploratory)	One-Sided or Two-Sided Test
1	Overall HPOG Impact	Confirmatory	Confirmatory	One-sided
1	Overall HPOG Impact	Secondary	Secondary	One-sided
1	Overall HPOG Impact	Exploratory	Exploratory	Two-sided
2	Overall HPOG Impact for Key Subpopulations	Confirmatory	Secondary	Two-sided
2	Overall HPOG Impact for Key Subpopulations	Secondary	Exploratory	Two-sided
2	Overall HPOG Impact for Additional Subpopulations	Confirmatory	Exploratory	Two-sided
3	Impact of Randomly Assigned Enhancement to HPOG	Confirmatory	Secondary	Two-sided
3	Impact of Randomly Assigned Enhancement to HPOG	Secondary	Exploratory	Two-sided
3	Variation in HPOG Impact Associated with Locally Adopted Program Components	Confirmatory	Exploratory	Two-sided
4	Variation in HPOG Impact Associated with Individual Program Participation	Confirmatory	Exploratory	Two-sided

5.3 Reporting Findings at Various Evidence Levels

For full reports of project results, we will include all of the outcomes and all subgroups and analyses detailed in this analysis plan. Discussion of each set of results will be accompanied by an appropriate explanation of the level of evidence that that set provides, along with the needed cautions associated with classes of analysis that increase risk of Type I error. In contrast, for reporting findings in very short (e.g., one-page) summaries common practice dictates that only confirmatory findings be reported, though we will extend this study's coverage to include secondary hypotheses as well. Given the prioritization of HPOG-Impact's Research Questions—its emphasis on the "what works" questions in addition to examining overall program impact—we plan to include all the test results for the *confirmatory* outcome in the 15-month analysis, for evidence that is either at the confirmatory or secondary levels. This means that we will include results from the experimental tests of HPOG program enhancements as well as the overall experimental test of HPOG's impact in short summaries. This decision also permits including overall impacts for key subgroups, should they be of interest to ACF.

5.4 Definitions of Model Terms

For reference, Exhibit 5.2 presents the definitions of the terms included in the analytic models presented in Sections 5.5 through 5.7. Where possible, we note what the letters/abbreviations stand for, and it is our

hope that these conventions make it easier for other analysts to understand the elements of the equations that follow.

Exhibit 5.2: Definitions of Model Terms

Name	Definition
Outcome and Covariates	
Y_{kji}	The outcome measure for individual *i* from division *j* and program *k*
T_{kji}	The standard HPOG program treatment indicator (1 for those individuals assigned to the standard HPOG treatment; 0 for the control group individuals; this is labelled "T" for "treatment")
E_{kji}	The enhanced HPOG program treatment group indicator (1 for only those individuals assigned to the enhanced HPOG treatment group; 0 otherwise; this is labelled "E" for "enhanced" treatment)
TE_{kji}	The HPOG program treatment group indicator (1 for those individuals assigned to the standard HPOG treatment or enhanced HPOG treatment groups; 0 for the control group individuals; this is labelled "TE" for the combination of standard "treatment" and "enhanced" treatment groups)
IC_{ckji}	Individual baseline characteristic *c* for individual *i* from division *j* and program *k* (grand mean centered), *c* = 1, . . ., *C* (this is labelled "IC" for "individual characteristics")
I_{gkj}	implementation feature *g* for division *j* in program *k* (grand mean centered), *g* = 1, . . ., *G* (these are labelled "I" for "implementation")
$\bar{I}_{gk}$	Implementation feature *g* averaged across divisions within program *k* (grand mean centered)
PC_{dkj}	Participant composition variable *d* for division *j* in program *k* (grand mean centered), *d* = 1, . . ., *D*; this is a division-level aggregation of the individual characteristics (ICs) (these are labelled "PC" for "participant composition")
P_{mk}	Program component *m* for program *k* (grand mean centered), *m* = 1, . . ., *M* including the experimentally varied enhancement components P_{Sk}, P_{Ak}, and P_{Ik} (these are labelled "P" for "program")
LC_{qk}	Local context variable *q* for program *k* (grand mean centered), *q* = 1, . . ., *Q* (these are labelled "LC" for "local context")
F_k	Omitted program-level factor F_k (which could be the aggregation to the program level of an omitted division-level factor) that influences treatment impact magnitudes (this is labelled "F" for "factor")
Model Coefficients	
α_{kj} (alpha)	The control group mean outcome (counterfactual) in division *j*
α_k	The control group mean outcome (counterfactual) in program *k*
α_0	The grand mean control group outcome
β_{kj} (beta)	The conditional impact of being offered the standard HPOG program for each division *j*
β_k	The conditional impact of being offered the standard HPOG program for each program *k*
β_0	The grand mean impact of the standard HPOG Treatment
δ_c (delta)	The effect of individual characteristic *c* on the mean outcome, *c* = 1, . . ., *C*
γ_c (gamma)	The influence of individual characteristic *c* on impact magnitude, *c* = 1, . . ., *C*
π_{ekj} (pi)	The impact of being offered an enhanced HPOG program that includes component *e* relative to the standard HPOG program for each division; this and the other subscripted "pi"s are program component impacts
π_{ek}	The impact of being offered the enhanced HPOG program, inclusive of component *e*, rather than the standard HPOG program without e, for each program
π_e	The grand mean impact of being offered the enhanced HPOG program inclusive of component *e*, rather than the standard HPOG program without *e*
π_m	The influence of program component *m* on impact magnitude, *m* = 1, . . ., *M*

Name	Definition
ζ_q (zeta)	The influence of local context variable *q* on impact magnitude, *q* = 1, . . ., *Q*
κ_q (kappa)	The effect of local context variable *q* on control group mean outcome
φ_g (phi)	The influence of implementation feature *g* on impact magnitude, *g* = 1, . . ., *G*
τ_g (tau)	The influence of participant composition variable *d* on impact magnitude, *d* = 1, . . ., *D*
λ (lambda)	The amount by which a one-unit change in F (an omitted confounder) alters impact magnitude
Error Terms	
ε_{kji} (epsilon)	A random component of the outcome for each individual
v_{kj}	A random component of control group mean outcome for each division
v_k	A random component of control group mean outcome for each program
u_{kj}	A random component of the standard program impact for each division
u_k	A random component of the standard program impact for each program
ω_{kj} (omega)	A random component of the enhanced program's incremental impact for each division
ω_k	A random component of the enhanced program's incremental impact for each program

5.5 Method for Estimating HPOG's Impact

In this subsection, we describe the analysis plan for addressing Research Question 1: What impact do the HPOG programs as a group have on outcomes for participants and their families? For each of the outcomes (confirmatory, secondary and exploratory) described in Section 2.2, we will estimate Intent to Treat (ITT) impacts of *being given access* to the basic HPOG program using a multi-level regression model that adjusts the difference between average outcomes for treatment and control group members by controlling for exogenous characteristics measured at baseline. Because policy makers usually can only *offer* access to a program, as is the case with HPOG, we compute the ITT estimate. Analyses described later (see Section 7) explore the implications of actually *participating* in selected program components. That model, described below, will be estimated using the combined sample of all individuals randomly assigned to the standard HPOG treatment group or the control group across the 23 grantees that the impact analysis will analyze.

5.5.1 Model Specification

We plan to estimate a three-level model.[18,19] The unit of analysis for level one is the individual sample member; the unit of analysis for level two is the division; and the unit of analysis for level three is the

[18] We note that the Evaluation Design Report presented a two-level model. As the project has evolved, recognition of the distinctions between and data availability at the program and division levels also revealed that a three-level model would be the more appropriate choice. The Evaluation Design Report's specification remains accurate but more general than what we now detail in this technical supplement.

[19] If we are unable to estimate a three-level model (e.g., if the model fails to converge or if division-level measures are excluded due to degree of freedom constraints), then we will consider collapsing to a two-level model for the analyses described in this subsection. Using a two-level model requires additional assumptions (e.g., homogenous impacts across higher-level units), but has the benefit of lesser computational demands than a three-level model.

program.[20,21] We plan to use a model with similar structural components for all impact analyses described in Sections 5.5 through 5.7, though model details such as the sample used and covariates included will vary across analyses. Section 5.4, above, provides a table that summarizes, in one place, the notation used throughout these models.

The level one regression equation depicted by Equation (5-1) models the relationship between an individual program participant's outcome *Y* and an HPOG program treatment indicator *T* while controlling for individual characteristics *IC*. To do this, the equation includes parameters for the conditional control group mean (α_{kj}) (i.e., the counterfactual, after adjusting for baseline covariates) and the treatment impact (β_{kj}) for each division *j* and program *k*. These parameters provide the dependent variables for level two of the model, as depicted in Equations (5-2) and (5-3).

Level One: Individuals

$$Y_{kji} = \alpha_{kj} + \beta_{kj}T_{kji} + \sum_c \delta_c IC_{ckji} + \varepsilon_{kji} \qquad \text{(eq. 5-1)}$$

where:

Y_{kji} = the outcome measure for individual *i* from division *j* and program *k*;

α_{kj} = the control group mean outcome (counterfactual) in division *j* (this is "alpha");

β_{kj} = the conditional impact of being offered the standard HPOG program for each division *j*;

T_{kji} = the standard HPOG program treatment indicator (1 for those individuals assigned to the standard HPOG treatment; 0 for the control group individuals);

δ_c = the effect of individual characteristic *c* on the mean outcome, *c* = 1, . . ., *C*;

IC_{ckji} = individual baseline characteristic *c* for individual *i* from division *j* and program *k* (grand mean centered), *c* = 1, . . ., *C*;[22] and

---

[20] An administrative division is a program intake location or locations with a dedicated case management and/or counseling staff that advises participants, connects them to education and training services, and provides participants with support services or refers them to these services. An administrative division may be a single intake location or may be multiple locations served by a single set of case managers and program administrators. Administrative divisions will be formed by combining such locations. Programs may have one or more such divisions. Fourteen HPOG programs have two or more divisions within them.

[21] A program is a unique set of services, training courses and personnel. Many grantees fund and operate one program; some fund multiple programs.

[22] We plan to grand mean center all individual characteristics so the values for β_j represent the treatment impact for the typical member of the full study sample (i.e., the sample member with mean values for all individual characteristics). As described by Hofmann and Gavin (1998), raw metric and grand mean centering options provide equivalent models. However, Kreft, De Leeuw, and Aiken (1995) recommends the use of grand mean centering instead of raw metric approaches because it usually results in a reduction of the covariance between the intercepts and slopes, thereby reducing potential problems associated with multicollinearity. Related, Hofmann and Gavin (1998) note that, in most all cases, group mean centering will produce models that are not equivalent to either raw metric or grand mean centering approaches. Though all three centering options are not equivalent, Kreft, De Leeuw, and Aiken (1995) conclude that "there is no statistically correct choice" among the three models. The choice between grand mean centering and group mean centering must be determined by theory. Bloom, Hill, and Riccio (2001) elect to grand mean center all independent variables, which allows them to interpret all slope and intercept coefficients as representing the typical individual from the typical site. We follow Bloom, Hill, and Riccio (2001) in scaling independent variables to be grand mean centered.

ε_{kji} = a random component of the outcome for each individual (this is "epsilon").

Level Two: Divisions

$$\beta_{kj} = \beta_k + u_{kj} \quad \text{(eq. 5-2)}$$

where,

β_k = the conditional impact of being offered the standard HPOG program for each program *k*; and

u_{kj} = a random component of the standard program impact for each division.

Additionally, we have:

$$\alpha_{kj} = \alpha_k + v_{kj} \quad \text{(eq. 5-3)}$$

where,

α_k = the control group mean outcome (counterfactual) in program *k*; and

v_{kj} = a random component of control group mean outcome for each division.

Level Three: Programs

$$\beta_k = \beta_0 + u_k \quad \text{(eq. 5-4)}$$

where,

β_0 = the grand mean impact of the standard HPOG Treatment; and

u_k = a random component of the standard program impact for each program.

Additionally, we have:

$$\alpha_k = \alpha_0 + v_k \quad \text{(eq. 5-5)}$$

α_0 = the grand mean control group outcome; and

v_k = a random component of control group mean outcome for each program.

Finally, we can simplify the above three-level model by substituting Equations (5-2) through (5-5) into Equation (5-1), which produces the following model:

$$Y_{kji} = \alpha_0 + \beta_0 T_{kji} + \sum_c \delta_c IC_{ckji} + \{\varepsilon_{kji} + v_k + v_{kj} + u_k T_{kji} + u_{kj} T_{kji}\} \quad \text{(eq. 5-6)}$$

The coefficient β_0 is the primary coefficient of interest because it equals the average impact of being offered standard HPOG relative to the counterfactual condition of no HPOG. We plan to use maximum likelihood procedures (which assume joint normal distributions for the random components) to estimate the above model.

We note further that this is a basic impact estimation model, with treatment indicator and individual baseline variables to control for the slight, inevitable, random variations between treatment and control group characteristics; its only uncommon feature is the addition of added error terms to account for the hierarchical structure of the data, which nests individuals within divisions within programs. Later, we add level-two and level-three characteristics to this model as we further explore the relative effectiveness of selected implementation features and program components.

5.6 Sensitivity Analysis

The Analysis Plan does not detail all possible sensitivity tests that the evaluation might undertake, but there is one in particular that we can report on at this time. Specifically, we anticipate subsetting the sample to conduct a sensitivity test of the influence of control group "contrast" in the evaluation. That is, at the outset of the study, we recognized that some of the HPOG grantees' programs were not markedly different from what was available in the community, with the ease of service access and additional structural support that HPOG offered being the main difference in these "low" contrast programs. While the 15-month follow-up survey will be the main source for information on the contrast in training and program experiences between treatment and control group members, we plan to demarcate at the outset the subset of study sites where expected Treatment-Control contrasts are not negligible: those in which the HPOG treatment is noticeably different from "business as usual." Some few, selected programs might be classified as being low-contrast. If these lower contrast programs also have smaller impacts, then they may be suppressing the impacts estimated across all programs; excluding them in a sensitivity analysis will allow us to judge the extent to which this is the case, and interpret the balance of results accordingly.

5.7 Method of Estimating Subgroup Impacts

Next, in this subsection, we describe the analysis plan for answering Research Question 2: To what extent do impacts vary across selected subpopulations?

To answer this question, we plan to use the impact model described in Section 5.5 while including a level-one interaction between the treatment indicator and the subgroup categories of interest. The coefficient on this interaction term will provide an ITT estimate of the impact of the HPOG program on the subgroup of interest and will be used to address questions regarding HPOG's impact on selected subgroups. We expect to analyze the impacts for several selected (exogenous) subgroups defined by baseline characteristics. Furthermore, we plan to test for impacts separately by subgroup and to test for differences between subgroups. We recognize that the sample sizes required for detecting differences in subgroup impacts are much larger than those required to detect the subgroup impacts themselves. It is likely that these tests will require a series of caveats that warn the reader against making too much of differences (or non-differences) in impacts.

The subgroup analysis will focus on personal characteristics identified in the logic model as interacting with the program components, program outputs and outcomes. Among the subpopulations of interest, we identify two kinds of subgroups. The first are those whose characteristics associate with a policy trigger or action. These as those which distinguish among individuals in the following categories, all measured at the point of random assignment:

- Education (five subgroups): attainment is less than 12th grade, a High School Equivalency (GED), some college but no degree, postsecondary degree but not a bachelor's and bachelor's degree or higher
- Barriers to Employment (four subgroups defined based on the number of barriers—childcare arrangements, transportation, an illness or health condition or alcohol or drug use—that fairly or very often interfere with school, work, job search or family responsibilities): no barriers, one barrier, two barriers, and three or more barriers
- Public Assistance (three subgroups, not mutually exclusive): receiving welfare, receiving WIC or SNAP, or not receiving public assistance

- Employment (two subgroups): employed and not employed
- Participation expectations (two subgroups): expect to participate in HPOG full-time or part-time

A central purpose of the subgroup analysis is to identify avenues for program improvement, and we believe the information provided from the analyses of subgroups listed above has a direct connection to that. The remaining subpopulations of interest are defined on the basis of the following constructs, which are not manipulable by policy, and are therefore analyzed for only descriptive purposes:

- Age (two subgroups): younger than 25 years old (typical post-secondary age) and 25 years old or older
- Race/Ethnicity (three subgroups): Hispanic, Non-Hispanic White, and Non-Hispanic Black
- Dependent Children (two subgroups): No dependent children and one or more dependent children

These measures are fully operationalized in Appendix A, Exhibits A.1–A.8.

Using the confirmatory outcome, we will investigate the impact of the HPOG program for all subpopulations of interest. For secondary outcomes, we will investigate the impact of the program for key subpopulations only. As described in Exhibit 5.1 above, the impact of HPOG on the confirmatory outcome for key subpopulations is designated a secondary hypothesis. The remaining subgroup analyses are exploratory.

5.8 Table Shells for Reporting Findings

Exhibit 5.3 is a sample table shell for reporting estimates of the overall HPOG impact and the impact of HPOG on specified subgroups (Sections 5.5 and 5.7, above). Columns (1) and (2) present regression-adjusted mean outcomes for the treatment and control groups, respectively. Columns (3) and (4) present the impact estimate, in both absolute and relative (to the control group mean) terms. Column (5) presents impact estimate standard error.

Exhibit 5.3: Sample Table Shell for In-Text Results—Regression-Adjusted Mean Outcomes and Impacts of HPOG Program

	Treatment Group Mean (1)	Control Group Mean (2)	Impact (3)	Percent Impact (4)	Standard Error (5)
Outcome Domain1					
Measure1	A	B	A-B*	%	
FILL IN…					
Outcome Domain2					
Mesure2					

Notes:
* Statistically significant, p<0.01
** Statistically significant, p<0.05
*** Statistically significant, p<0.10

In addition to this core information, we anticipate appendix tables that provide additional information. Specifically, the appendix tables (a sample of which appears as Exhibit 5.4) will also include a designation of the level of evidence associated with the reported test as confirmatory (C), secondary (S), or exploratory (E). Appendix tables will also present the sample size, a minimum detectable effect, or

smallest true impact with an 80 percent chance of significance for that outcome, and the R-squared, which denotes the proportion of outcome variation explained by the analysis model.[23]

Referring to the minimum detectable effect (MDE) results, we note that MDEs are typically calculated and reported as part of study design. Indeed, in Section 4.6 of the Evaluation Design Report, we included MDEs based on *assumptions* about the standard deviation of the outcomes, the explanatory power of covariates, and other details. When a study fails to reject the null-hypothesis of zero impact, readers and researchers alike are left to wonder whether the intervention had no impact or if the intervention had an impact too small to be detected. We propose to revisit the calculation of MDEs during impact analysis to provide additional context for null-findings because we believe this will be useful for judging the circumstances in which some impacts are not flagged as statistically significant. Specifically, we propose to calculate and present MDEs for each hypothesis test. The MDEs may be used in interpreting the findings and will be included to help the reader understand the size of impact the study was powered to detect. These MDEs will differ from those presented in the Evaluation Design Report, because they will be calculated from the *actual* variance structure of the analytic model which reflects observed outcome and covariate data.

Following Schochet (2008b), the MDE formula can be expressed as follows:

$$MDE = Factor(\sigma, \beta, df) * SE(impact),$$

where $SE(impact)$ is the standard error of the impact estimate and $Factor(\sigma, \beta, df)$ is a constant that is a function of the significance level (α), statistical power (β), and the number of degrees of freedom (*df*) .[24] For these after-the-fact MDE analyses, we will use a statistical significance level of 0.05 and power of 80 percent.

[23] There are multiple R-squared coefficients that matter in multi-level models, one for each level. We will include all salient R-squared coefficients together in that column, with a slash between them and a footnote to the column headers explaining the meaning of each entry (e.g., individual-level R-squared/division-level R-squared).

[24] $Factor(\sigma, \beta, df)$ can be expressed as $[T^{-1}(\alpha) + T^{-1}(\beta)]$ for a one-tailed test and $[T^{-1}(\alpha/2) + T^{-1}(\beta)]$ for a two-tailed test, where $T^{-1}(.)$ is the inverse of the student's t distribution function with *df* degrees of freedom (Schochet, 2008b).

Exhibit 5.4: Sample Table Shell for Detailed Results in Appendix—Regression-Adjusted Mean Outcomes and Impacts of HPOG Program

	Level of Evidence (1)	Treatment Group Mean (2)	Control Group Mean (3)	Impact (4)	Percent Impact (5)	Signifi-cance Level (6)	Standard Error (7)	Minimum Detectable Effect (8)	Sample Size (9)	R-Squared (10)
Outcome Domain1										
Measure1	C	A	B	A-B	(A-B)/B	**			N	
FILL IN…										
Outcome Domain2										
Motivation										
FILL IN…										

Notes: Level of evidence denotes whether the impact hypothesis is confirmatory (C), secondary (S), or exploratory (E).
* Statistically significant, p<0.01
** Statistically significant, p<0.05
*** Statistically significant, p<0.10

6. The Influence of Program Components on Average Impacts

PLEASE NOTE: The material in this Chapter is superseded by an Amendment that is located here: https://www.acf.hhs.gov/sites/default/files/opre/hpog_impact_analysis_plan_rq3_revised_plans_finalv4_508.pdf [PDF]

In this section, we discuss our plan for answering Research Question 3: Which locally adopted program components influence average impacts?

In Section 6.1, we describe our plan for estimating the impact of selected program components. Randomly assigning individuals to two variants of the program within designated program sites allows for the comparison of different programmatic scenarios: one in which the program adopts the component of interest, and an alternative where the program does not adopt the component of interest.[25] The contrast in outcomes between these two treatment groups reveals the contribution of the component as an add-on to the main program. That is, it shows the difference in impact that adding the component as an enhancement causes, given the already-existing features of the program. This is the best information for deciding whether to include the selected component as part of the standard program model going forward.

In Section 6.2, we discuss our plan to exploit division-level variation in implementation features and program-level variation in program components to estimate the relationship between these intervention features and impact magnitude. This analysis is conducted for all confirmatory outcomes, as described in Section 2.2 above. In contrast to the analyses mentioned above, these analyses are non-experimental in nature because they rely on non-randomly occurring variation across programs and divisions. In Section 6.3, we describe plans to use the unbiased, purely experimental estimate of the impact of the enhancement feature from the three-arm random assignment sites as a benchmark for selecting non-experimental estimates with low bias. Exhibit 5.2 provides a table that summarizes, in one place, the notation used throughout these models; it may be helpful for the reader to have this exhibit's details nearby to promote understanding of the model terms used in this section.

6.1 Analysis of Randomly Assigned Program Enhancements

In this section, we describe the plan to test experimentally the impact of three promising program enhancements on all confirmatory and secondary outcomes: facilitated peer support groups, emergency assistance, and non-cash incentives. HPOG staff and program participants in programs with a strong peer support component have noted that the support and associated accountability is considered to be one of the most important program elements. Program staff cite unanticipated financial need as a major reason for program dropout, and believe that easier access to emergency funds could buffer participants in times of crisis and improve program retention and completion. Non-cash incentives may also lead to improved participant outcomes by motivating desirable in- and out-of program behaviors. For example, in a job retention and advancement program, a results-based incentive might reward those individuals who stay employed for six months, while a behavior-based program might reward individuals who achieve perfect attendance.

Exhibit 6.1 shows the grantees in which enhancement components were randomly assigned, with expected sample sizes. In the exhibit, "NV" refers to "natural variation," indicating whether the

[25] In comparison, the analyses described in Section 7 attempt to uncover the extent to which individuals' participation in various HPOG components leads to differential impacts on individual trainees.

designated program component existed naturally within the grantee before it was added as a randomized-to enhancement in some places.

Exhibit 6.1: Grantee with Experimental Tests of Enhancement Components, by Type

HPOG-Impact Grantee	Peer Support	Emergency Assistance	Non-Cash Incentives
Bergen Community College		(9 Programs)[a] T=490 TE=359	(Essex CC)[b] T=195 TE=55
Eastern Gateway Community College	NV	NV	NV
Kansas Department of Commerce			
Schenectady County Community College		NV	NV
New Hampshire Office of Minority Health	T=257 TE=218	NV	
Milwaukee Area WIB			
South Carolina Department of Social Services			T=201 TE=127
Buffalo and Erie County WDC	T= 358 TE=60	NV	
Gateway Community and Technical College (KY)		NV	T=118 TE=65
Central Community College			NV
Suffolk County Department of Labor		NV	T=262 TE=92
Pensacola State College		NV	
WIB SDA-83 Inc. (LA)			
Research Foundation of CUNY-Hostos Comm. Coll.		T=276 TE=196	
Will County WIB		(some)[c] NV	
Full Employment Council	NV	T=144 TE=122	NV
Central Susquehanna Intermediate Unit		NV	
The WorkPlace	T=159 TE=113	NV	
Alamo Comm. Coll. District and Univ. Health System			T= 115 TE=62
Edmonds Community College		NV	
HPOG/PACE Grantee	**Peer Support**	**Emergency Assistance**	**Non-Cash Incentives**
Pima County Community College District			
San Diego Workforce Partnership		NV	
Workforce Dev. Council of Seattle-King County		NV	

Source: HPOG-Impact Evaluation Design Implementation Plans.
Notes: Black cells indicate that a sufficient contrast exists and the grantee is implementing the enhancement for an experimental test of its effectiveness. Gray cells indicate that there is not sufficient contrast ("NV" indicates that these programs might be used to explore the natural variation that exists on this program component). White cells indicate that sufficient contrast exists for such a test, but the grantee is not implementing an enhancement.
[a] Nine HPOG programs within the Bergen Community College grantee are implementing the enhancement.
[b] The Essex Community College program within the Bergen Community College grantee is implementing the enhancement.
[c] There is not a sufficient contrast at a subset of the grantee's programs.

Our ability to obtain statistically significant findings showing that an enhancement component affects impact magnitude, when in fact it does, will be limited by available sample sizes—particularly by the number of programs and divisions that randomly assigned cases to both enhanced and standard programs (plus a control group). While the number of individuals included in the groups, as shown in Exhibit 6.1, may be adequate to support reasonable power at that level, limitations at other levels exist for each of the enhancement components, as follows:

- 7 divisions within 3 programs implemented three-arm random assignment with facilitated peer support as their enhancement;
- 15 divisions within 11 programs implemented three-arm random assignment with emergency assistance as their enhancement; and
- 10 divisions within 5 programs implemented three-arm random assignment with non-cash incentives as their enhancement.

Section 6.4, below, explains how non-experimental evidence on the contribution of these three program components can be combined with experimental evidence to increase the study's ability to statistically detect true non-zero contributions of the enhancement components.

6.1.1 Model Specification

For the experimental analysis of the effect of the selected component, we will use a three-level model to estimate program impacts controlling for program and individual factors. The unit of analysis for level one is the individual sample member; the unit of analysis for level two is the division; and the unit of analysis in level three is the program. Compared to the three-level model in Section 5.5, here we include an added incremental impact term for the randomly assigned to enhancement feature.

The level one regression equation depicted by Equation (6-1) below uses data from individuals to model the relationship between an outcome *Y* and an overall HPOG treatment indicator (which denotes whether the participant was assigned to either the standard HPOG treatment or enhanced HPOG treatment) and an enhanced treatment indicator while controlling for individual characteristics. The equation also includes the conditional control group mean and treatment impact for each division. The conditional impact estimates (β_{kj} and π_{ekj}) and control group means (α_{kj}) for each program provide the dependent variables for level two of the model, as depicted in Equations (6-2), (6-3), and (6-4). The π_{ekj} parameter varies with the sample of sites being analyzed: sites with random assignment to facilitated peer support ($e = S$), sites with random assignment to emergency assistance ($e = A$), and sites with random assignment to non-cash incentives ($e = I$). For reference, Exhibit 5.2 presents the definitions of the terms included in the models presented next, and we do not restate those definitions here for parsimony.

Level One: Individuals

$$Y_{kji} = \alpha_{kj} + \beta_{kj}TE_{kji} + \pi_{ekj}E_{kji} + \sum_c \delta_c IC_{ckji} + \varepsilon_{kji} \quad \text{(eq. 6-1)}$$

Level Two: Divisions

$$\beta_{kj} = \beta_k + u_{kj} \quad \text{(eq. 6-2)}$$

$$\pi_{ekj} = \pi_{ek} + \omega_{kj} \quad \text{(eq. 6-3)}$$

and:

$$\alpha_{kj} = \alpha_k + v_{kj} \quad \text{(eq. 6-4)}$$

Level Three: Programs

$$\beta_k = \beta_0 + u_k \quad \text{(eq. 6-5)}$$

$$\pi_{ek} = \pi_e + \omega_k \quad \text{(eq. 6-6)}$$

and:

$$\alpha_k = \alpha_0 + v_k \quad \text{(eq. 6-7)}$$

We can simplify the above three-level model by substituting Equations (6-2) through (6-7) into Equation (6-1), which produces the following model:

$$Y_{kji} = \alpha_0 + \beta_0 TE_{kji} + \pi_e E_{kji} + \sum_c \delta_c IC_{ckji} + \{\varepsilon_{kji} + v_k + v_{kj} + u_k TE_{kji} + u_{kj} TE_{kji} + \omega_k E_{kji} + \omega_{kj} E_{kji}\}, \quad \text{(eq. 6-8)}$$

Here, π_e is the primary coefficient of interest: it provides an estimate of the impact of being offered the enhanced HPOG program relative to the standard HPOG program. Conducting the analysis separately for facilitated peer support enhancement programs, emergency assistance enhancement programs; and non-

cash incentives enhancement programs provides the experimental estimates of the contribution of those program components to the overall impact magnitude, the various π_e terms for the selected enhancements. We define these estimates as follows:

- $\hat{\pi}_S^X$ provides an experimental estimate of π_e when data from programs that randomly assign to facilitated peer support are analyzed;
- $\hat{\pi}_A^X$ provides an experimental estimate of π_e when data from programs that randomly assign to emergency assistance are analyzed; and
- $\hat{\pi}_I^X$ provides an estimate of π_e when data from sites that randomly assign to non-cash incentives are analyzed.

We plan to use maximum likelihood procedures (which assume joint normal distributions for the error terms) to estimate the above model.

6.2 Examining the Role of Non-Randomized Program Characteristics

In addressing Research Question 3, our goal is to understand how program characteristics influence the magnitude of intervention impacts so that stronger program designs can be developed and adopted in the future. The programs that randomize to three experimental arms provide the best evidence on these questions but only for samples of limited size and only for the three HPOG components being tested experimentally as program enhancements. The study will also take advantage of the naturally occurring variation in the specific services offered by programs (program components) and in how these services are delivered (implementation features) across research sites to extend findings about how these intervention features may influence impacts.

This section describes our analytic approach to estimating the influence of division-level implementation features and participant composition measures, as well as program-level component and local context measures on impact magnitude. This nested structure allows us to observe variation in impact magnitudes across divisions and programs. We can then explore how this variation relates to observed variation in program components, implementation features, participant composition and local context measures.

We are particularly interested in how program components and implementation features relate to impact magnitudes, as these characteristics could be incorporated into future programs. The extent to which impact varies by participant composition and local context is important for understanding how, when and for whom the program works. While these measures are not manipulable by program managers, they may be of policy relevance in terms of program targeting.

6.2.1 Model Specification

To relate program characteristics to impact magnitude, we extend the multi-level model in Section 5.5 by interacting the treatment indicator with measures of program characteristics. We will conduct this analysis on confirmatory outcomes—educational progress for the First Follow-Up report and educational progress and earnings for the Intermediate Outcomes Report. We will use an empirical approach to select division, program- and local context-level measures to be included in the impact analysis models. We plan to estimate the model using the combined sample of all individuals in the standard HPOG treatment group

or in the control group across all 23 grantees that are part of the main impact analysis.[26,27] In this three-level model the unit of analysis at level one is the individual sample member; the unit of analysis at level two is the division; and the unit of analysis at level three is the program.

Level 1: Individuals

$$Y_{kji} = \alpha_{kj} + \beta_{kj} T_{kji} + \sum_c \delta_c IC_{ckji} + \varepsilon_{kji} \quad \text{(eq. 6-9)}$$

Level 2: Divisions

$$\beta_{kj} = \beta_k + \sum_g \varphi_g I_{gkj} + \sum_d \tau_d PC_{dkj} + u_{kj} \quad \text{(eq. 6-10)}$$

and:

$$\alpha_{kj} = \alpha_k + v_{kj} \quad \text{(eq. 6-11)}$$

Level 3: Programs

$$\beta_k = \beta_0 + \sum_m \pi_m P_{mk} + \sum_q \zeta_q LC_{qk} + u_k \quad \text{(eq. 6-12)}$$

and:

$$\alpha_k = \alpha_0 + \sum_q \kappa_q LC_{qk} + v_k \quad \text{(eq. 6-13)}$$

Combining the elements of the above three-level model produces the following:

$$Y_{kji} = \alpha_0 + \beta_0 T_{kji} + \sum_c \delta_c IC_{ckji} + \sum_q \kappa_q LC_{qk} + \sum_m \pi_m P_{mk} T_{kji} + \sum_q \zeta_q LC_{qk} T_{kji} + \sum_g \varphi_g I_{gkj} T_{kji} + \sum_d \tau_d PC_{dkj} T_{kji} + \{\varepsilon_{kji} + v_k + v_{kj} + u_k T_{kji} + u_{kj} T_{kji}\} \quad \text{(eq. 6-14)}$$

In equation 6-14, the local context measures (LC_{qk}), program components (P_{mk}), implementation features (I_{gkj}) and participant composition measures (PC_{dkj}) are all multiplied by the treatment indicator. These interaction terms capture the influence of the measure on impact magnitude. In addition, the local context measures enter the model directly, capturing the influence of the economic environment on control group outcomes. This specification does not allow participant composition measures to affect control group outcomes, assuming that the individual-level characteristics (IC_{ckji}) included in the model are more salient to individual outcomes. Measures of program components and implementation features are not allowed to affect control group outcomes, as control group members did not access these services.

The extent to which program components and implementation features relate to impact magnitudes is our primary interest. In the equations above, the coefficients π_m for $m = 1,\ldots,M$ capture the relationship between impact magnitude and program components. The coefficients φ_g for $g = 1, \ldots, G$, the relationship between implementation features and impact magnitudes. Because estimates of these parameters are identified by the natural variation in program components and implementation features within the HPOG program, these estimates are non-experimental and will be interpreted accordingly. That is, the estimates of the relative effects of these aggregate-level variables are associated with the characteristics of programs and grantees that choose to put them into place (only individuals and not program components, for example, are randomized).

[26] Depending on the measures empirically selected for inclusion in the model, we may exclude individuals who were randomly assigned to an enhanced treatment group.

[27] We plan to use maximum likelihood procedures (which assume joint normal distributions for the error terms) to estimate the above model.

6.2.2 Degrees of Freedom and Measure Selection

This analysis focuses on measures defined at the division-, program- and local context-level, and therefore the sample sizes at each of these levels determine the number of these covariates we can enter into the model. Overall, the sample includes 26 local context areas, 42 programs and 87 divisions.[28] A standard guideline requires five observations per covariate. Based on this, if we only include division-level measures, we could include at most about 17 measures. However, if we include program-level measures as well, we need to account for the degrees of freedom at each level. Exhibit 6.2 below lists possible combinations of the number of measures at the local context-, program- and division-level that guarantee at least five observations at the appropriate level per measure.[29]

Exhibit 6.2: Number of Local Context-, Program- and Division-Level Covariates

Local Context	Program	Division
2	1	8
2	2	5
2	3	3
1	1	11
1	2	9
1	3	7
1	4	5
1	5	2
1	6	0
0	3	10
0	4	8
0	5	6
0	6	4
0	7	2
0	8	0

In response to the limited number of measures we can include at each of these levels, we use a two-pronged selection approach. First, we identified a list of candidate measures based on our expectations regarding their relationship to the effectiveness of the program.[30] Second, we will use an empirical approach to select among these measures to fully specify the model. Described in much greater detail below, the empirical specification approach seeks to reduce bias in the non-experimental estimates. One potential draw-back of an empirical specification approach is that the method might select measures that are not of primary interest to the field or are not readily interpretable. However, the first step in which we limit the candidates for inclusion to a short list of policy-relevant measures guarantees that the model selected by the algorithm will be interpretable and interesting to the field.

[28] Local context measures can be thought of as approximately grantee-level. In equations 6-12 and 6-14 above, the local context variables enter the model at the program-level. However, observations for programs within the same grantee have the same value of these measures, because these programs experience the same economic environment. Appendix A provides more detail on these local context measures are obtained for programs and grantees from Census and BLS data sources.

[29] We exclude individual-level covariates from discussion of degrees of freedom intentionally. Because our individual sample size is so large, including the roughly thirty baseline covariates described in Section 2.2 does not affect our degrees of freedom at the local context-, program- or division-level.

[30] Section 4.1 discusses our hypotheses regarding these relationships and provides citations to the literature.

6.3 Reducing Attributional Bias

We will use within-program experimental evidence to reduce the bias in cross-site estimates of the contributions of enhancement components to program impacts by specifying local-, program- and division-level measures for inclusion in the model. To do so, we will use information about the unbiased, purely experimental estimates of the impact of the enhancement features from the three-arm random assignment sites as benchmarks for moving the corresponding non-experimental estimates closer to unbiasedness. The method may also provide a way to increase the reliability of *all* non-experimental estimates, not just those for intervention features varied experimentally.[31] In what follows, we describe how bias arises when estimating the influence of program components and implementation features on impact magnitude in Section 6.2; the plan for finding the least-biased Equation (6-14) model specification; and a description of the programs and divisions used to conduct the bias reduction strategy.

The notation in this section is simplified from the previous section. Although the approach described here is used to select local context measures, program components, implementation features and participant composition measures for inclusion in the model, this section is framed in terms of selecting among program components for ease of exposition.

Throughout this section, we will use $\hat{\pi}_e^N$ to refer to the non-experimental estimate of the influence of one of the randomly assigned enhancement components—peer support, emergency assistance or non-cash incentives—on the magnitude of the impact. Here, the superscript *N* stands for non-experimental, and the subscript *e* stands for enhancement. The non-experimental estimate comes from the multi-level model described in Section 6.2. In contrast, we will use $\hat{\pi}_e^X$ to refer to the experimental estimate of the impact of the same enhancement, where the superscript *X* stands for experimental and the estimate comes from the model presented in Section 6.1. The non-experimental estimate of the influence of a generic program component is denoted $\hat{\pi}_m^N$.

6.3.1 Source and Measurement of the Bias Threat

The bias concern arises from the possibility that one or more division-level factor or program-level factor that influences treatment impact magnitudes has been omitted from the model in Section 6.2 and hence does not appear in Equation (6-14). For example, dynamism of executive leadership in the HPOG service delivery organization may affect impact magnitude at the program level—and be unmeasurable—or peer effects of participating in HPOG as part of a group of highly motivated students may boost impacts at the division level but not be feasible to enter into the model. Appendix D provides a version of the analysis model that omits important determinants of impact of this sort, and demonstrates how such omitted factors create bias in the $\hat{\pi}_m^N$ estimates produced by the model.

Fortunately it is possible to address and reduce this bias in the $\hat{\pi}_e^N$ estimates empirically. The technique hinges on using $\hat{\pi}_e^X$, estimated using random assignment by Equation (6-8), as an unbiased benchmark against which to judge $\hat{\pi}_e^N$ using the following formula:

$$\text{estimated bias}(\hat{\pi}_e^N) = |\hat{\pi}_e^N - \hat{\pi}_e^X|. \qquad \text{(eq. 6-15)}$$

We describe in the next subsection our plan for finding the least-biased estimates of $\hat{\pi}_e^N$ by strategically choosing the Equation (6-14) model specification.

[31] Work is in progress to establish whether this is the case. Intuitively it seems plausible that this is the case, if omitted factors that led some grantees to agree to adopt the three enhancement features also play a role in grantee decisions to adopt certain other program components and implementation features in their main HPOG programs.

6.3.2 Choice of Facilitated Peer Support as the Benchmark

We will use facilitated peer support as the enhancement component for which we measure bias to inform model specification. We choose to rely on a single measure rather than measure bias for all three enhancements based on degree of freedom concerns. If we used all three enhancements to measure bias, we would need to allocate three program-level measures in Equation 6-14 to those three enhancements. Given the degree of freedom constraints discussed above, we prefer to retain these degrees of freedom for measures selected to reduce bias.

From among the three enhancements, we focus attention on facilitated peer support because we believe the sites that adopted the component as an experimental enhancement made decisions about inclusion of this component similarly to the grantees elsewhere in the sample that adopted facilitated peer support on their own (i.e., through what we have referred to as "natural variation"). In contrast, the other enhancements appear to have been selected through a distinct process that might not yield insight into the selection bias affecting program characteristics that reflect natural variation. Without knowledge of how the information would factor into the impact analysis, project staff involved in recruiting grantees to implement an enhancement classified each grantee that implemented an enhancement component into one of three categories: at the ends of the spectrum were "early adopters" or "late adopters" of the enhancement, and in the middle were some that "required some time" to choose to adopt the enhancement. All of the peer support enhancement programs fell into the "early adopters" category, as such making them more like the grantees that started their HPOG programs with peer support in place. In contrast, all of the programs that implemented non-cash incentives as an enhancement "required some time" to do so or were "late adopters." As such, they are probably less like the average grantee that naturally offers a non-cash incentive program. Finally, programs that implemented emergency assistance split between the "early" and "late" categories. Based on this classification, we believe that programs that chose to implement either emergency assistance or noncash incentives as part of their standard program probably differed from the average grantee in their approach to implementing these components.

However, concerns that naturally occurring instances of facilitated peer support differ systematically from facilitated peer support delivered as a randomly assigned enhancement to the standard program remain valid. Therefore, we plan to restrict the sample for the model specification exercise to the set of grantees, programs and divisions that either offer facilitated peer support as a randomly assigned enhancement to HPOG or do not offer facilitated peer support. By excluding naturally occurring instances of facilitated peer support from the model specification, we guarantee that the non-experimental impact and the experimental impact are measuring the same program component and that in expectation differences between these quantities are due to bias. We can then specify the model to reduce the bias in estimates of the impact of facilitated peer support identified by variation in adoption of the enhancement component. If the process of adopting the enhancement component is influenced by the same omitted variables as the process of selecting program components and implementation features to include in the standard program, then the model that reduces this bias may also reduce bias in the estimates identified by natural variation in program characteristics.

Exhibit 6.3 below presents the numbers of grantees, programs and divisions that offer each enhancement as part of standard HPOG, as a randomly assigned enhancement and do not offer the enhancement. The sample used for modeling focuses on the 73 divisions, 40 programs and 21 grantees that either offer facilitated peer support as an enhancement or do not offer it. The number of covariates we can include in the model is determined by the number of divisions, programs, and grantees available for the analysis. The fact that the number of divisions, programs, and grantees in the sample used for the analysis described in this section is similar to the number of divisions, programs, and grantees in the full sample

(as displayed in Exhibit 6.1) is another reason to prefer peer support as our benchmark case. If we focused on either of the other two enhancements, we would lose a much larger proportion of the sample, as both emergency assistance and noncash incentives were more widely offered as part of the standard HPOG program, in comparison to facilitated peer support.

Exhibit 6.3: Number of Grantees, Programs and Divisions Offering Each Enhancement

	Offered in Standard HPOG	Offered as an Enhancement	Not Offered	Total
Peer Support				
Grantees	2	3	18	23
Programs	2	3	37	42
Divisions	13	7	67	87
Emergency Assistance				
Grantees	13[b]	3[a]	9[a,b]	23
Programs	16	11	15	42
Divisions	32	14	41	87
Noncash Incentives				
Grantees	4	5[a]	15[a]	23
Programs	4	5	33	42
Divisions	20	10	57	87

Notes:
[a] Bergen Community College appears twice in the grantee row, because it has programs that offer the component as an enhancement and programs that do not offer the component.
[b] Will County WIB appears twice in the grantee row, because it offers the component as part of standard HPOG at some programs and does not offer the component at other programs.

6.3.3 Finding the Least-Biased Model Specification

The estimate of bias in Equation (6-15) can be used to adjust the specification of the Equation (6-14) model in Section 6.2 to include more of the division-, program- and local context-level determinants of impact magnitude, or to include them in a more accurate functional form, to see if the measured bias becomes smaller.

The procedure for finding the least-biased specification of Equation (6-14) involves constructing an estimate of bias for a specific set of local context measures, program components, implementation features and participant composition measures by following three steps:

Step 1: Obtain an experimental estimate of the impact of the facilitated peer support enhancement ($\hat{\pi}_S^X$) by estimating the Equation 6-8 for the sample of divisions that provide facilitated peer support as an experimental enhancement to individuals randomized to the second treatment arm.

Step 2: Derive a non-experimental estimate of the influence of the facilitated peer support component on impact size ($\hat{\pi}_S^N$) by estimating Equation 6-14 for the sample of divisions that either offer facilitated peer support as an randomly assigned enhancement to HPOG or do not offer facilitated peer support. In the sites that offer peer support as a randomly assigned enhancement, include individuals assigned to the enhanced treatment group and to the control group, omitting individuals assigned to the standard HPOG treatment group.

Step 3: Construct a measure of the estimated bias:

$$\text{estimated bias}(\hat{\pi}_S^N) = \left|\hat{\pi}_S^X - \hat{\pi}_S^N\right| \qquad \text{(eq. 6-16)}$$

To find the least-biased model specification, these three steps are repeated, systematically varying the specification of Equation (6-14) to depict how program components, implementation features, participant

composition measures, and local context variables affect impact magnitude. A detailed plan for systematically altering Equation (6-14) in this way is described in Appendix C, which includes a stopping rule based on the number of degrees of freedom used up in expanding the model to take account of potentially more and more determinants of impact.

6.3.4 Summary

In sum, we have unbiased, purely experimental estimate of the impact of the facilitated peer support enhancement, $\hat{\pi}_S^X$, from the three-arm random assignment sites and will use this estimate as benchmark for moving the corresponding non-experimental estimates, $\hat{\pi}_S^N$, closer to unbiasedness. Doing so may also improve the non-experimental methodology for attributing program-level impacts to program components generally for all the cross-site estimates, $\hat{\pi}_1^N, \ldots, \hat{\pi}_e^N, \ldots, . \hat{\pi}_M^N$.

Because this approach has not been employed previously, we will test this approach prior to applying it to the HPOG-Impact analysis. If Monte Carlo analyses show that the method does not improve the identification of program components and implementation features that drive impact under reasonable assumptions, we will use a theory-driven approach to specifying Equation 6-14.

6.4 Table Shells for Reporting Findings

To report the experimental impacts of HPOG program enhancements (Section 6.1), a table similar to the one presented in Exhibit 5.3 will be used.

Exhibit 6.4 provides a sample table shell for reporting estimates of the contribution of non-randomized program components to impact magnitude (Section 6.2). Exhibit 6.4 also presents standard errors of the impact estimates, and the sample size of individuals and clusters.

Exhibit 6.4: Sample Table Shell—Estimates of the Contribution of Non-Randomized Program Components to Impact Magnitude

	Impact (1)	Standard Error (2)	Significance Level (3)
Program Components			
Program Component 1			
Program Component 2			
Implementation Features			
Implementation Feature 1			
Implementation Feature 2			
Local Context Measures			
Local Context Measure 1			
Local Context Measure 2			
Participant Composition Measures			
Participant Composition Measure 1			
Participant Composition Measure 2			
Sample Size			
Individuals			
Divisions			
Programs			

Notes:
* Statistically significant, p<0.01
** Statistically significant, p<0.05
*** Statistically significant, p<0.10

7. Exploiting Variation in Individual-level Participation in Program Components

In this section, we discuss our plan for answering Research Question 4: To what extent does participation in a particular HPOG component (or components) change the impact on individual trainees? This question focuses on the individual's experience of program services. A key goal of this research question is to improve program design and inform sharper service delivery.

HPOG-Impact will not only capitalize on the cross-location planned and natural variation (see Section 6), but it will also capitalize on the substantial individual-level variation in program experiences. The impact estimates produced by the analyses described in Section 6 concern how much program enhancements contribute to impact magnitudes and how effective unenhanced programs, with all their components, are for the average participant. We can also examine the role of individual's participation in selected program components and/or achievement of intermediate milestones in response to Research Question 4. This involves an analysis of "endogenous" subgroups that can be addressed by the "ASPES"—or analysis of symmetrically predicted endogenous subgroups—as established in the literature. We refer readers to other published documents to learn the details of the methodological approach to be employed and here focus on the details of the specific HPOG application.[32] In brief, this kind of analysis focuses on treatment-induced experiences (mediators) that can be analyzed as if they are exogenous traits (moderators), thereby capitalizing on the experimental evaluation design. The results of this line of analysis will yield impacts for subgroups that are defined along the lines of their program experiences or their intermediate outcomes as identified in the program's logic model and elaborated further in this section.

Exhibit 7.1 lists the specific endogenous subgroups of interest to HPOG-Impact. We classify the subgroups into two types: individual mediators and program-related mediators. Achievement of program outputs or milestones can be thought of as individual mediators. These include earning a professional credential or license or an academic certificate or degree. The program-related mediators can be any of those *program* characteristics that this research focuses on, including participation in any of the many program components described in Section 4. We analyze the effect of programs' *offering* these components by the method described in Section 6; and this analysis of endogenous subgroups will consider the effect of individual *participation* in each given program component. Additional information on these measures, including their data sources, appears in Section 4.2.

[32] Peck (2003) first described the approach, and Peck (2013) revisits it in the context of the subsequent decade of evaluation research. Harvill, Peck and Bell (2013) and Abadie, Chingos, and West (2014) consider how to carry out subgroup identification without introducing bias. Bell and Peck (2013) further consider the method's assumptions. Moulton, Peck, and Bell (2014) is the Social Impact Policy Pathfinder (SPI-Path) | Individual User Guide, which details the steps and analytic decisions along the way with sample SAS and Stata code for executing ASPES. Applications that might be especially useful for understanding how ASPES works in practice include Peck and Bell (2014), which considers the case of Head Start quality, and Moulton, Peck, and Dillman (2014), which considers the case of neighborhood quality in the Moving to Opportunity Experiment.

Exhibit 7.1: Selected Endogenous Subgroups (Mediators of Impact Magnitude)

Subgroup	Part of Logic Model	Outcomes of Interest	Sample Restrictions
Obtained a professional, state or industry license or credential	Program outputs/Short run outcomes	Employment, Earnings	Combined sample of all individuals across the 23 grantees
Completed a degree (AA, BA or higher)	Program outputs/Short run outcomes	Employment, Earnings	Combined sample of all individuals across the 23 grantees
Participated in peer support	Program content/design	Educational Progress	Restrict sample to programs that offered the service to at least some individuals TE vs. C in enhancement programs T vs. C in natural variation programs
Received emergency assistance	Program content/design	Educational Progress	Restrict sample to programs that offered the service to at least some individuals TE vs. C in enhancement programs T vs. C in natural variation programs
Received a non-cash incentive	Program content/design	Educational Progress	Restrict sample to programs that offered the service to at least some individuals TE vs. C in enhancement programs T vs. C in natural variation programs
Used personal counseling services	Program components	Educational Progress	Restrict sample to programs that offered the service to at least some individuals
Used academic assistance services (tutoring, etc.)	Program components	Educational Progress	Restrict sample to programs that offered the service to at least some individuals

As noted in Exhibit 7.1, the analysis will consider three outcomes: educational progress, employment and earnings, each of which is fully defined in Section 2.

This line of analysis capitalizes on variation across individuals to consider which program components contribute most to HPOG impacts or achievement of intermediate milestones that are important in individuals' subsequent employment and earnings trajectories. We focus on two program milestones (earning as credential/license or degree) and five program components as detailed in Exhibit 7.1. We consider what the employment and earnings impacts are for those who achieve selected intermediate milestones; and we consider how participation in the selected program components influences educational progress. The program's logic model implies that certain program design elements (e.g., making counseling and other services available) are intended to help individuals stay connected to the program and complete it, achieving desired milestones, which then associate with more favorable labor market outcomes.

We choose to supplement impact analyses using ASPES for the three program components (peer support, emergency assistance, non-cash incentives) that are offered in some places through a three-armed experiment and two others (personal counseling, academic tutoring/counseling) that were identified by HPOG Impact sites as key to participant success and were offered in most places. We judge this added analysis to enrich what the study can learn about the relative influences of these intermediate variables, with specific focus on *participation* in a selected program component. This is in contrast to earlier analyses where we discuss how to estimate the effects of gaining *access to* (or being *offered*) the selected program component.

Consider peer support. Two of the HPOG-Impact grantees have always offered peer support as part of their program; and an additional three grantees added the option in year three, rationing access to peer support via a lottery. In the ASPES-based work, we will pool all five of those programs, and use baseline

characteristics to help us understand which individuals *took up* the offer of peer support, and then analyze the impact of the program for them, relative to those who did not participate in any peer support (despite that it might have been available to them). In addition to providing estimates for the role of peer support in generating program impacts, along the way this analysis will reveal which sorts of people took up the focal program components.

7.1 Table Shells for Reporting Findings

To report findings generated using the ASPES method, we will present impacts on the selected endogenous subgroups of interest, as shown in Exhibit 7.2. In addition, we will include the impact estimates for predicted subgroups, which are used as a mechanism for identifying the impacts on actual subgroups, in appendix tables. Reporting impact estimates for predicted subgroups is not essential to interpreting the impacts on actual subgroups. However, reporting impacts on predicted subgroups can provide transparency in that the reader then can both better understand the steps of the ASPES method. Exhibit 7.3 shows how we will present impacts on predicted subgroup members, in an appendix to the main document.

Exhibit 7.2: Sample Table Shell—Impact on Endogenous Subgroups

	Impact on Actual Subgroup A				Impact on Actual Subgroup B				
	Mean Outcome for Subgroup A Control Group	Impact on Subgroup A (β)	Standard Error of Impact on Subgroup A	Obs.	Mean Outcome for Subgroup B Control Group	Impact on Subgroup B (β)	Standard Error of Impact on Subgroup B	Obs.	Sig. Diff. in Subgroup Impacts: (2) - (6)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Outcome 1	.	.	.	.	.	.	.	.	***
Outcome 2	.	.	.	.	.	.	.	.	
Outcome 3	.	.	.	.	.	.	.	.	

Notes: Impact estimates converted to represent actual Subgroup A and Subgroup B members. We compute control group mean outcomes for Subgroups A and B by subtracting actual subgroup impact estimates from calculated mean outcomes for members of the subgroup observed in the treatment group. This is necessary since we cannot directly observe which control group members would have joined which a given subgroup had they been offered the treatment.

*** statistically significant, p<0.01

** statistically significant, p<0.05

* statistically significant, p<0.10

Exhibit 7.3: Sample Appendix Table Shell—Impact on Predicted Subgroups

	Impact on Predicted Subgroup A				Impact on Predicted Subgroup B				
	Mean Outcome for Subgroup A Control Group	Impact on Subgroup A (β)	Standard Error of Impact on Subgroup A	Obs.	Mean Outcome for Subgroup B Control Group	Impact on Subgroup B (β)	Standard Error of Impact on Subgroup B	Obs.	Sig. Diff. in Subgroup Impacts: (2) - (6)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Outcome 1	.	.	.	.	.	.	.	.	***
Outcome 2	.	.	.	.	.	.	.	.	
Outcome 3	.	.	.	.	.	.	.	.	

Notes:

*** statistically significant, p<0.01

** statistically significant, p<0.05

* statistically significant, p<0.10

8. HPOG Impact Study Schedule and Deliverables

This final section lists, in Exhibit 8.1, the timeline for the Impact Study's remaining activities and deliverables.

Exhibit 8.1: Remaining Project Tasks, Deliverables, and Timeline

Task	Subtask/Description	Date(s)
Impact Study	Sample intake (random assignment, baseline data collection)	Mar. 2013 – Nov. 2014
	15-month follow-up survey data collection	Jun. 2014 – Sep. 2016
	36-month follow-up survey data collection[a]	Mar. 2016 – Apr. 2018
Reporting and Dissemination	Report to OCSE on findings using NDNH data[b]	Jan. 2016 – Apr. 2019
	Draft and final Analysis Plan, including sample baseline characteristics	Sep. 2014 – Aug. 2015
	Draft and final impact study report based on 15-month follow-up survey	May 2017 – Sep. 2017
	Draft and Final Special Topics Paper: TANF Recipients in the Health Profession Opportunity Grants (HPOG) Program	Mar. 2015 - Aug. 2015
	Draft and Final Special Topics Paper: Methods for Understanding Impact Variation	Jan. 2016 – Jun. 2016
	Draft and Final Special Topics Paper (topic TBD)	by Sep. 2017
	Draft and final impact study report based on 36-month follow-up survey[a]	Dec. 2018 – Apr. 2019
Data Files and Documentation	Submit data files and documentation to ACF	Sep. 2017

Notes:
[a] This data collection and associated reports will be conducted under the CPIO Study.
[b] After the impact study report based on the 15-month follow-up survey is completed, these annual reports will be conducted under the CPIO Study.

References

Abadie, A., Chingos, M. M., & West, M. R. (2014). *Endogenous stratification in randomized experiments* (Harvard University Working Paper). Cambridge, MA: Harvard University. Retrieved from http://www.ksg.harvard.edu/fs/aabadie/stratification.pdf

Bell, S. H., & Peck, L. R. (2013). Using symmetric predication of endogenous subgroups for causal inferences about program effects under robust assumptions: Part two of a method note in three parts. *American Journal of Evaluation, 34*(3), 413–426. doi: 10.1177/1098214013489338

Bloom, H. S., Hill, C. J., & Riccio, J. (2001). *Modeling the performance of welfare-to-work programs: The effects of program management and services, economic environment, and client characteristics* (Working paper). New York: Manpower Demonstration Research Corporation.

Bloom, H. S., Hill, C. J., & Riccio, J. A. (2003). Linking program implementation and effectiveness: Lessons from a pooled sample of welfare-to-work experiments. *Journal of Policy Analysis and Management, 22*(4), 551–575.

Bos, J., Huston, A., Granger, R., Duncan, G., Brock, T., & McLoyd, V. (1999). *New hope for people with low incomes: Two-year results of a program to reduce poverty and reform welfare.* New York: Manpower Demonstration Research Corporation.

Brodkin, E. Z. (1997). Inside the welfare contract: Discretion and accountability in state welfare administration. *Social Service Review, 71,* 1–33.

Education Longitudinal Study of 2002 (ELS:2002). (2002.) Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Endel, B., Anderson, N., & Kelley, J. (2011). *Achieving ambitious goals: Case studies of scaling-up programs for advancing low-skilled adults.* Boston, MA: Jobs for the Future.

Estrada, R. A. (2010). *How to build bridge programs that fit into a career pathway.* Chicago: Instituto del Progreso Latino.

Evans, T. (2010). *Professional discretion in welfare services: Beyond street-level bureaucracy.* London: Ashgate.

Fein, D. J. (2012). *Career pathways as a framework for program design and evaluation: A working paper from the Innovative Strategies for Increasing Self-Sufficiency (ISIS) Project* (OPRE Report No. 2012-30). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Geckeler, C. (2008). *Helping community college students cope with financial emergencies: Lessons from the Dreamkeepers and Angel Fund Emergency financial aid programs.* New York: Manpower Demonstration Research Corporation.

Grant-Vallone, E., Reid, K., Umali, C., & Pohlert, E. (2004). An analysis of the effects of self-esteem, social support, and participation in student support services on students' adjustment and commitment to college. *Journal of College Student Retention: Research, Theory & Practice, 5*(3), 255–274.

Harvill, E. L., Peck, L. R., & Bell, S. H. (2013). On overfitting in analysis of symmetrically predicted endogenous subgroups from randomized experimental samples: Part three of a method note in three parts. *American Journal of Evaluation, 34*(4), 545–566. doi: 10.1177/1098214013503201

Hinckley, R., & Hull, D. (2009). Excerpt from Adult Career Pathways: Providing a second chance in public education. *NCPN Connections*, April (special issue).

Hofmann, D. A., & Gavin, M. B. (1998). Centering decisions in hierarchical linear models: Implications for research in organizations. *Journal of Management, 24,* 623–641.

Hsueh, J., Alderson, D. P., Lundquist, E., Michalopoulos, C., Gubits, D., & Fein, D. (2012). *The Supporting Healthy Marriage Evaluation: Early impacts on low-income families, technical supplement* (OPRE Report 2012-27). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Institute of Education Sciences. (2014). What Works Clearinghouse. What Works Clearinghouse procedures and standards handbook version 3.0. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/reference_resources/wwc_procedures_v3_0_draft_standards_handbook.pdf.

Izrael, D., Battaglia, M. P., & Frankel, M. R. (2009). Extreme survey weight adjustment as a component of sample balancing (aka raking) (Paper 247-2009). Presented at SAS Global Forum 2009. Retrieved from http://support.sas.com/resources/papers/proceedings09/247-2009.pdf

Jobs for the Future. (2010). *The breaking through practice guide.* Boston: Jobs for the Future.

Karp, M. M. (2011). *Toward a new understanding of non-academic student support: Four mechanisms encouraging positive student outcomes in community college* (Working Paper 28). New York: Community College Research Center, Teachers College, Columbia University.

Kazid, R., & Liebowitz, M. (2003). *Changing courses: Instructional innovations that help low-income students succeed in community college.* New York: Manpower Demonstration Research Corporation.

Kreft, I. G. G., De Leeuw, J., & Aiken, L. S. 1995. The effect of different forms of centering in Hierarchical Linear Models. *Multivariate Behavioral Research, 30,* 1–21.

Moulton, S., Peck, L. R., & Bell, S. H. (2014). *Social Policy Impact Pathfinder (SPI-Path) Analytic Suite: SPI-Path|Individual User Guide.* Bethesda, MD: Abt Associates.

Moulton, S., Peck, L. R., & Dillman, K. (2014). Moving to Opportunity's impact on health and well-being among high dosage participants. *Housing Policy Debate*, *24*(2), 415–446. doi: 10.1080/10511482.2013.875051.

Peck, L. R. (2003). Subgroup analysis in social experiments: Measuring program impacts based on post treatment choice. *American Journal of Evaluation, 24*(2), 157–187. DOI: 10.1016/S1098-2140(03)00031-6

Peck, L. R. (2013). On analysis of symmetrically-predicted endogenous subgroups: Part one of a method note in three parts. *American Journal of Evaluation, 34*(2): 225-236. DOI: 10.1177/1098214013481666

Peck, L. R., & Bell, S. H. (2014). *The role of program quality in determining Head Start's impact on child development* (OPRE Report 2014-10). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Pusser, B., & Levin, J. (2009). *Re-imagining community colleges in the 21st century: A student-centered approach to higher education.* Washington, DC: Center for American Progress.

Richburg-Hayes, L., Brock, T., LeBlanc, L., Paxson, C., Rouse, C. E., & Barrow, L. (2009). *Rewarding persistence: Effects of a performance-based scholarship program for low-income parents.* New York: Manpower Demonstration Research Corporation.

Schochet, P. Z. (2008a). *Technical methods report: Guidelines for multiple testing in impact evaluations* (NCEE 2008-4018). National Center for Education Evaluation and Regional Assistance.

Schochet, P. Z. (2008b). *Technical methods report: Statistical power for regression discontinuity designs in education evaluations* (NCEE 2008-4026). National Center for Education Evaluation and Regional Assistance.

Schwarzer, R., & Jerusalem, M. (1995). Generalized self-efficacy scale. In J. Weinman, S. Wright, & M. Johnston, *Measures in health psychology: A user's portfolio* (pp. 35–37). Windsor, UK: NFER-Nelson. Retrieved from http://userpage.fu-berlin.de/health/engscal.htm

Scrivener, S., & Weis, M. J. (2009). *More guidance, better results? Three-year effects of enhanced student services program at two community colleges.* New York: Manpower Demonstration Research Corporation.

Soares, L. (2010). *Community college and industry partnerships.* Paper presented at the White House Summit on Community Colleges, Washington, DC.

Stephens, R. P. (2009). *Charting a path: An exploration of the statewide career pathway efforts in Arkansas, Kentucky, Oregon, Washington and Wisconsin.* Seattle, WA: Seattle Jobs Initiative.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures for student attrition* (2nd ed.). Chicago: University of Chicago Press.

Werner, A., Dun Rappaport, C., Stuart, J. B., & Lewis, J. (2013). *Literature review: Career pathways programs* (OPRE Report #2013-24). Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data.* Cambridge, MA: The MIT Press.

Appendix A. Measures' Operationalization Details

This appendix provides detailed operationalization of the measures to be used in the analysis. It includes baseline measures (be they covariates, descriptive variables, subgroup-defining variables), all outcomes, division- and program-level variables (including components, implementation features, local contextual variables), and individual-level participation measures. These measures were introduced in the body of the analysis plan. This appendix adds information on the specific variable or items to be used to construct the measure and how we plan to combine those data to create the analysis variable.

Exhibit A.1: Operationalization of Baseline Demographic Measures

Construct	Variable Description	Operationalization/Details	Designated Use(s)	Data Source(s) Items
Sex	Respondent is male (binary)	Binary variable takes on value - 0 if female - 1 if male Missing otherwise (Omit largest category from model)	Descriptor Covariate	HPOG-Impact site: PRS New Participant Form Gender HPOG/PACE site: BIF B07_sex
Marital Status	Marital status (categorical)	Categorical measure capturing marital status: - Married - Separated or Divorced - Widowed - Never Married - Missing Substantive categories match those from which the respondent selected. Follow definitions from HPOG annual report to identify and define the missing category.	Descriptor	HPOG-Impact site: PRS Intake Form marital HPOG/PACE site: BIF B08_marstat
Dependent Children	Number of dependent children (categorical)	Categorical description of the number of dependent children: - None - One or two - Three or more - Missing Note: In PACE sites, this measure captures the number of children living in the home for whom the participant is the guardian	Descriptor	HPOG-Impact site: PRS Intake Form dependent_children HPOG/PACE site: BIF B15a_num_chn B15b_num_guardian
Dependent Children	Parent to one or more dependent children (binary)	Binary variable that takes on value: - 1 if parent to one or more dependent children - 0 if have 0 dependent children - Missing otherwise Note: In PACE sites, this measure captures the number of children living in the home for whom the participant is the guardian	Covariate Subgroup Identifier	HPOG-Impact site: PRS Intake Form dependent_children HPOG/PACE site: BIF B15a_num_chn B15b_num_guardian

Construct	Variable Description	Operationalization/Details	Designated Use(s)	Data Source(s) Items
Race/Ethnicity	Ethnicity (Hispanic/Latino) and race (series of mutually exclusive binary variables)	Series of mutually exclusive binary variables: - Hispanic/Latino - Non-Hispanic White/Caucasian - Non-Hispanic Black/African-American - Non-Hispanic Asian, Native Hawaiian, or Pacific Islander - Non-Hispanic American Indian or Native Alaskan - Non-Hispanic Two or more races (Omit largest category from models) Follow definitions from HPOG annual report.	Descriptor Covariate Subgroup Identifier	HPOG-Impact site: PRS Intake Form race_w race_a race_hpi race_b race_na ethnicity HPOG/PACE site: BIF B09_Hispanic B10_race_white B10_race_black B10_race_Am_Indian B10_race_Pacific B10_race_Asian
Age	Age (categorical)	Categorical measure capturing: Less than 20 Years - 20-24 years - 25-29 years - 30-34 years - 35-39 years - 40-49 years - 50+ years - Missing Categorical measure created from continuous measure of age at random assignment	Descriptor	HPOG-Impact site: PRS New Participant DOB create_date hpog_program_reg HPOG/PACE site: BIF B02_dob R_RA_Date_Assigned
Age	Age (continuous)	Two variables capturing a continuous measure of age: - Age in years at random assignment - Age in years at random assignment squared 1. Create a date of random assignment. Initially set date of random assignment to HPOG_Program_Reg. If HPOG_Program_Reg is missing, set date of random assignment to Create_date 2. Calculate age at random assignment from date of birth and date of random assignment. Set age at random assignment to missing if date of birth is missing. 3. Create squared age from age at random assignment	Covariate	HPOG-Impact site: PRS New Participant DOB create_date hpog_program_reg HPOG/PACE site: BIF B02_dob R_RA_Date_Assigned

Construct	Variable Description	Operationalization/Details	Designated Use(s)	Data Source(s) Items
Age	Typical postsecondary age (younger than 25 years old at random assignment) (binary)	Binary variable created from the age in years at random assignment: Value 1 if less than 25 at random assignment Value 0 if 25 or older at random assignment Missing if age at random assignment is missing	Subgroup Identifier	HPOG-Impact site: PRS New Participant DOB create_date hpog_program_reg HPOG/PACE site: BIF B02_dob R_RA_Date_Assigned
Born Outside U.S.	Born outside the U.S. (binary)	Binary variable that takes on value: 0 if born in the United States, Puerto Rico, Guam, the U.S. Virgin Islands, or Northern Mariana 1 if born elsewhere Missing otherwise* Note: In PACE sites, this measure captures birth in one of the 50 U.S. states or Washington, D.C.	Descriptor Covariate	HPOG-Impact site: PRS Intake Form citizenship HPOG/PACE site: BIF B11_born_USA

Exhibit A.2: Operationalization of Baseline Education Background Measures

Construct	Variable Description	Operationalization/Details	Designated Use(s)	Data Source(s) Items
Education	Completed education measured in years (categorical)	Categorical variable that captures years of education completed: - Less than 12th Grade - High School Equivalency/GED - High School Graduate - 1-3 Years of College/Technical School - 4 Years or More of College - Missing Note: In PACE sites, individuals who report "1+ years of college no degree" or "Associates Degree" are included in the "1-3 Years of College/Technical School" category. Individuals who report "Bachelor's or higher" are included in the "4 Years or More of College" category.	Descriptor	HPOG-Impact site: PRS Intake Education Form highest_education HPOG/PACE site: BIF B17_educ
Education	Educational attainment measured by degrees and credentials completed (series of mutually exclusive binary variables)	Series of mutually exclusive binary variables that capture educational attainment: - Less than 12th grade - High School equivalency (GED) - High School Graduate - Some college, no degree - Received a post-secondary degree, not a Bachelor's - Received Bachelor's degree or higher	Subgroup Identifier	HPOG-Impact site: PRS Intake Education Form highest_education degree_received_hs degree_received_ged degree_received_other_postsec degree_received_ba HPOG/PACE site: BIF B17_educ
Education	Attained postsecondary degree prior to random assignment (binary)	Binary variable that takes on a value of 1 if any of the following are true: * Received a BA degree * Received another postsecondary degree (option on PRS in each HPOG-Impact site) * Received an AA degree (option on BIF in HPOG-PACE sites) 0 if the respondent indicated that he or she received neither type of degree Missing otherwise	Descriptor; Pre-intervention measure of outcome Covariate	HPOG-Impact site: PRS Intake Education Form degree_received_other_postsec degree_received_ba HPOG/PACE site: BIF B17_educ
Education	Occupational Skills License or Certification prior to random assignment (binary)	Binary variable that takes on a value of 1 if Occupational Skills License or Certification prior to random assignment 0 if no Occupational Skills License or Certification prior to random assignment Missing otherwise	Descriptor; Pre-intervention measure of outcome Covariate	HPOG-Impact site: PRS Intake Education Form DEGREE_RECEIVED_CERTIFICATE HPOG/PACE site: BIF B18_voctech_cert

Construct	Variable Description	Operationalization/Details	Designated Use(s)	Data Source(s) Items
Education	Completed license, certification or degree prior to random assignment (binary)	Binary variable that takes on a value of 1 if any of the following are true: * received a BA degree * received another postsecondary degree (option on PRS in each HPOG-Impact site) * received an AA degree (option on BIF in HPOG-PACE sites) * received Occupational Skills License or Certification 0 if none of the above are true Missing otherwise	Pre-intervention measure of outcome Covariate	HPOG-Impact site: PRS Intake Form DEGREE_RECEIVED_BA DEGREE_RECEIVED_OTHER_POSTSEC DEGREE_RECEIVED_CERTIFICATE HPOG/PACE site: BIF B17_educ B18_voctech_cert
Basic Skills	Literacy assessed at 8th grade level or higher (binary)	Binary variable that takes on a value of 1 if participant literacy assessed at 8th grade level or higher, 0 if not, and missing otherwise Note: This measure is not available for PACE sites	Descriptor; Covariate	HPOG-Impact site: PRS Intake Form NUMERACY
Basic Skills	Numeracy assessed at 8th grade level or higher (binary)	Binary variable that takes on a value of 1 if participant numeracy assessed at 8th grade level or higher, 0 if not, and missing otherwise Note: This measure is not available for PACE sites	Descriptor; Covariate	HPOG-Impact site: PRS Intake Form LITERACY
Skills Course Attendance	Course attendance in - Adult Basic Education classes - English as a Second Language classes - Vocational, Technical or Trade School classes - Classes in how to succeed in school - Classes in how to succeed at work (series of binary variables)	Series of mutually exclusive binary variables that capture course attendance: - Adult Basic Education classes - English as a Second Language classes - Vocational, Technical or Trade School classes - Classes in how to succeed in school - Classes in how to succeed at work	Descriptor; Covariate	HPOG-Impact site: PRS Supplemental Baseline Form attend_abe attend_esl attend_voc_tech_trade attend_school_success attend_work_success HPOG/PACE site: BIF B19_attend_ABE B19_attend_ESL B19_attend_vocational B19_attend_school_succ B19_attend_work_succ

Construct	Variable Description	Operationalization/Details	Designated Use(s)	Data Source(s) Items
Educational Expectations	Level of education expected to complete, measured in degrees and credentials (categorical measure)	Level expected to complete - No additional - GED - High school diploma - Alternative non-academic credential (option on PRS in HPOG-Impact site) - Associate's degree - Bachelor's degree - Graduate degree - Not reported Follow definitions from HPOG annual report.	Descriptor	HPOG-Impact site: PRS Supplemental Baseline Form expected_highest_ed HPOG/PACE site: BIF B20_expect_educ

Exhibit A.3: Operationalization of Baseline Income and Benefits Measures

Construct	Variable Description	Operationalization/Details	Designated Use(s)	Data Source(s) Items
Earnings	Average quarterly wage received during the four quarters prior to the quarter of random assignment (continuous)	Average quarterly wage received during the four quarters prior to the quarter of random assignment (continuous)	Pre-intervention measure of outcome; Covariate	NDNH
Income	Household income over 12 months prior to random assignment (categorical)	Household income over 12 months prior to random assignment - $9,999 or Less - $10,000 to $14,999 - $15,000 to $19,999 - $20,000 to $29,999 - $30,000 to $39,999 - $40,000 or More - Missing Note: In PACE sites, this measure captures family income, rather than household income	Descriptor	HPOG-Impact site: PRS Intake Form HOUSEHOLD_INCOME HPOG/PACE site: BIF B27_tot_fam_income_cats b27_tot_fam_income
Income	Individual income over 12 months prior to random assignment (categorical)	Individual income over 12 months prior to random assignment - $0 - $1 to $9,999 - $10,000 to $14,999 - $15,000 to $19,999 - $20,000 to $29,999 - $30,000 or More - Missing Note: This measure is not available for PACE sites	Descriptor	HPOG-Impact site: PRS Intake Form INDIVIDUAL_INCOME
Public Assistance	Public assistance use, by source (series of binary variables)	Series of binary variables describing public assistance use - Welfare - WIC/SNAP Note: HPOG/PACE participants reported whether they received any form of welfare, whereas HPOG/Impact participants reported only TANF.	Descriptor; Pre-intervention measure of outcome; Covariate; Subgroup Identifier	HPOG-Impact site: PRS Intake Form TANF SNAP HPOG-Impact site: PRS Supplemental Baseline Form (WIC) ib_wic HPOG/PACE site: BIF B26b_WICSNAP B26c_PA

Exhibit A.4: Operationalization of Baseline Employment and Expected Time Use Measures

Construct	Variable Description	Operationalization/Details	Designated Use(s)	Data Source(s) Items
Employment	Proportion of quarters employed during the four quarters prior to the quarter of random assignment (continuous)	1. Create binary variables that indicate employment during each of the quarters 2. Average the quarters	Descriptor Pre-intervention measure of outcome of Outcome Covariate	NDNH
Employment	Currently employed (self-report) (binary)	Binary variable that takes on a value of 1 if he or she is currently employed (self-report), 0 if he or she is not, and missing otherwise	Subgroup Identifier	HPOG-Impact site: PRS Intake Form CURRENT_EMPLOYMENT_ENROLL HPOG/PACE site: BIF B24_curr_work
Employment in Healthcare	Ever employed in a healthcare job (binary)	Binary variable that takes on a value of 1 if he or she has ever worked in a health care profession/occupation, 0 if he or she has not, and missing otherwise Note: This measure is not available for PACE sites	Pre-intervention measure of outcome; Covariate	HPOG-Impact site: PRS Intake Form HC_WORK
Hours	Number of hours worked for last week of employment if currently employed (categorical)	Number of hours worked for last week of employment: - Not working - <20 hours per week - 20-34 hours per week - 35+ hours per week Create categorical measure from continuous measure of hours worked Note: information only requested from individuals who report current employment. Report “not working” if not currently employed.	Descriptor	HPOG-Impact site: PRS Intake Education Form WEEKLY_HOURS_CURRENT CURRENT_EMPLOYMENT_ENROLL HPOG/PACE site: BIF B24_curr_hours B24_curr_work
Hours Expectations	Expected working hours (categorical)	Expected Working Hours - Not working - <20 hours per week - 20-34 hours per week - 35+ hours per week Create categorical measure from continuous measure of hours expected	Descriptor	HPOG-Impact site: PRS Supplemental Baseline Form expected_hours HPOG/PACE site: SAQ S02a_Work_Hours

Construct	Variable Description	Operationalization/Details	Designated Use(s)	Data Source(s) Items
Employment Expectations	Expect to be working for pay in the next few months (binary)	Binary variable that takes on a value of 1 if he or she expects to be working for pay in the next few months, 0 if he or she is not, and missing otherwise	Covariate	HPOG-Impact site: PRS Supplemental Baseline Form expected_hours HPOG/PACE site: SAQ S02a_Work_Hours
Participation Expectation	Expect to participate in HPOG full-time or part-time (categorical)	Categorical measure capturing: Expected to participate in HPOG full-time Expected to participate in HPOG part-time Not reported	Descriptor Subgroup Identifier	HPOG-Impact site: PRS Supplemental Baseline Form expected_school_enroll HPOG/PACE site: SAQ S01_Future_School

Exhibit A.5: Operationalization of Baseline Life Challenges Measures

Construct	Variable Description	Operationalization/Details	Designated Use(s)	Data Source(s) Items
Limited English Proficiency	Limited English proficiency (binary)	Binary variable that takes on a value of 1 if he or she has limited English proficiency, 0 if he or she does not, and missing otherwise Note: Limited English Proficiency is defined as speaking English "not well" or "not at all" for HPOG/PACE sample members who speaks a language other than English at home	Descriptor; Covariate; Subgroup Identifier	HPOG-Impact site: PRS Intake Form LIMITED_ENGLISH HPOG/PACE site: BIF B12a_other_lang b12b_speak_english
Barriers to Employment	Barriers to school/work (% at fairly or very often), including child care arrangements, transportation, health (series of binary variables)	Barriers to school/work: - Child care arrangements - Transportation - An illness or health condition - Alcohol or drug use For each potential barrier, binary variable that takes on a value of 1 if barrier interferes with school, work, job search or family responsibilities: * Fairly Often * Very Often Value 0 if: * Never * Almost Never * Sometimes Missing otherwise	Descriptor	HPOG-Impact site: PRS Supplemental Baseline Form often_child_care often_transportation often_illness often_alcohol_drug HPOG/PACE site: SAQ S15a_Child_Care S15b_Transportation S15c_Alcohol S15d_Illness
Barriers to Employment	Number of barriers that fairly often or very often interfere with school, work, job search or family responsibilities (integer measure, transformed to series of binary variables to identify subgroups)	Series of mutually exclusive binary variables based on the number of barriers that fairly often or very often interfere with school, work, job search or family responsibilities: - No barriers - One barrier - Two barriers - Three or more barriers	Pre-intervention measure of outcome; Covariate; Subgroup Identifier	HPOG-Impact site: PRS Supplemental Baseline Form often_child_care often_transportation often_illness often_alcohol_drug HPOG/PACE site: SAQ S15a_Child_Care S15b_Transportation S15c_Alcohol S15d_Illness

Exhibit A.6: Operationalization of Outcomes at 15-Month Follow-Up

Domain	Variable Description	Operationalization/Details	Outcome Designation	Data Source(s)
Educational Progress	Completion of training or ongoing enrollment in training	Binary variable with value 1 if one or more of the following is true: * Completed a degree (Associates or higher) since RAD * Completed a professional, state, or industry certificate, license or credential since RAD * Currently taking classes for college credit * Currently receiving occupational training AND did not start any occupational training that was not completed Value 0 otherwise Missing if don't know/refused	Confirmatory Outcome	15-month follow-up survey 18, 22, 23, 25a, 25b Note: education/credential items text changed to measure completion since random assignment
Earnings	Wages received during the 5th quarter after quarter of random assignment	Continuous variable measuring wages received during the 5th quarter after quarter containing random assignment Value 0 if not employed	Secondary Outcome	NDNH
Employment	Employed during the 5th quarter after quarter of random assignment	Binary variable with value 1 if employed during the 5th quarter after quarter containing random assignment Value 0 otherwise	Secondary Outcome	NDNH
Employment in Healthcare	Currently employed in a healthcare job or (if unemployed) worked for pay at some point after random assignment and most recent job was in healthcare	Binary variable with value 1 if one or both of the following holds * Currently employed and is in healthcare * Currently not employed, worked for pay since RAD and most recent job was in healthcare Value 0 otherwise Missing if don't know/refused	Secondary Outcome	15-month follow-up survey Q36, Q36d, 37, 37d
Job Benefits	Current or most recent job offers health insurance	Binary variable with value 1 if one or both of the following holds * Currently employed and job offers health insurance * Currently not employed, worked for pay since RAD and most recent job offered health insurance Value 0 otherwise Missing if don't know/refused	Secondary Outcome	15-month follow-up survey Q36,Q36f,Q37,Q37f
Public Assistance Benefits	Individual receipt of cash public assistance (TANF) in the prior month	Binary variable that takes on a value of 1 if the individual reports receiving TANF, 0 otherwise	15-month follow-up survey	15-month follow-up survey Q51

Exhibit A.7: Operationalization of Outcomes at 36-Month Follow-Up

Domain	Variable Description	Operationalization/Details	Outcome Designation	Data Source(s)
Earnings	Wages received during the 12th quarter after quarter containing random assignment	Continuous variable measuring wages received during the 12th quarter after quarter containing random assignment Value 0 if not employed	Confirmatory Outcome	NDNH
Educational Progress	Completion of training	To be completed after the 36-month survey is finalized as part of the impact analysis plan for the 36-month report.	Confirmatory Outcome	15-month follow-up survey 36-month follow-up survey

Exhibit A.8: Operationalization of Additional Outcomes at 15-Month Follow-Up

Domain	Variable Description	Operationalization/Details	Outcome Designation	Data Source(s) Items
Barriers to Employment	Number of barriers that very often interfere with school, work, job search or family responsibilities	Integer values ranging from 0 to 4 1. Assign the following values to the items in Question 45 in 15-month follow-up survey 1 "Very Often" 0 "Sometimes" or "Never" Missing "Don't Know" or "Refused" 2. Sum over the barriers Justification for ignoring the "sometimes" category: it is normal to have things like childcare and transportation sometimes interfere with school, work, job search, or family responsibilities.	Exploratory Outcome	15-month follow-up survey Q45
Earnings	Cumulative wages received during the five quarters after the quarter containing random assignment	Continuous measure of wages (zero and positive values) 1. Infer zero wages for quarters with no wages reported 2. Sum over five quarters after the quarter containing random assignment	Exploratory Outcome	NDNH
Economic Status	Personal income received from all sources	Continuous measure reported verbatim. If individuals refuse/don't know and then provide a categorical response: assume the middle value of the category. Perform sensitivity checks to verify appropriateness of the assumption. Missing if don't know/refused on prompt for categorical response	Exploratory Outcome	15-month follow-up survey Q52
Economic Status	Household income received from all sources	Continuous measure reported verbatim. If individuals refuse/don't know and then provide a categorical response: assume the middle value of the category. Perform sensitivity checks to verify appropriateness of the assumption. Missing if don't know/refused on prompt for categorical response	Exploratory Outcome	15-month follow-up survey Q54
Economic Status	Used loans in parents name to pay for school or living expenses	Binary variable (1 for yes, 0 for no, missing otherwise)	Exploratory Outcome	15-month follow-up survey Q26
Economic Status	Used loans in either own name or parents name to pay for school or living expenses	Binary variable (1 for yes, 0 for no, missing otherwise)	Exploratory Outcome	15-month follow-up survey Q26

Domain	Variable Description	Operationalization/Details	Outcome Designation	Data Source(s) Items
Economic Status	Personally received ANY government assistance in the prior month	Binary variable that takes on a value of 1 if the individual reports receiving any one of the following types of assistance: * TANF * SNAP * WIC * UI * Medicaid * Subsidized Child Care * Section 8/Public Housing * LIHEAP * FRPL Value 0 otherwise Missing if don't know/refused	Exploratory Outcome	15-month follow-up survey Q51
Public Assistance Benefits	Number of major welfare programs (TANF, SNAP, Medicaid) from which the individual received benefits in the prior month	Integer values ranging from 0 to 3 1. Code each source (TANF, SNAP, Medicaid) as 1 "Yes", 0 "No", Missing "Don't Know" or "Refused" 2. Sum the indicators	Exploratory Outcome	15-month follow-up survey Q51
Public Assistance Benefits	Household received ANY government assistance in the prior month	Binary variable that takes on a value of 1 if the individual reports household members receiving any one of the following types of assistance: * TANF * SNAP * WIC * UI * Medicaid * Subsidized Child Care * Section 8/Public Housing * LIHEAP * FRPL Value 0 otherwise Missing if don't know/refused	Exploratory Outcome	15-month follow-up survey Q53
Public Assistance Benefits	Number of major welfare programs (TANF, SNAP, Medicaid) from which the household received benefits in the prior month	Integer values ranging from 0 to 3 1. Code each source (TANF, SNAP, Medicaid) as 1 "Yes", 0 "No", Missing "Don't Know" or "Refused" 2. Sum the indicators	Exploratory Outcome	15-month follow-up survey Q53

Domain	Variable Description	Operationalization/Details	Outcome Designation	Data Source(s) Items
Education	Completed a college degree (Associates, Bachelors or higher)	Binary variable with value 1 if one or both of the following is true: * Completed an Associate's degree since RAD * Completed a Bachelor's degree or above since RAD Value 0 otherwise Missing if don't know/refused Note: In HPOG-Impact data collection, education/credential items text changed to measure completion since random assignment	Exploratory Outcome	15-month follow-up survey Q25a
Education	Earned any college credits since random assignment	Binary variable (1 for yes, 0 for no, missing otherwise)	Exploratory Outcome	15-month follow-up survey Q16
Education	Obtained a professional, state or industry certificate, license or credential since random assignment	Binary variable with value 1 if individual reports: * Obtained a professional, state or industry certificate, license or credential Value 0 if individual reports either of the following: * No formal training * Some coursework or training but no professional, state or industry certificate, license or credential Missing if don't know/refused Note: In HPOG-Impact data collection, education/credential items text changed to measure completion since random assignment	Exploratory Outcome	15-month follow-up survey Q25b
Education	Completed a degree (AA, BA or higher) or obtained a credential (professional, state or industry certificate, license or credential) since random assignment	Binary variable with value 1 if one or more of the following is true: * Completed an Associate's degree since RAD * Completed a Bachelor's degree or above since RAD * Obtained a professional, state or industry certificate, license or credential Value 0 otherwise Missing if don't know/refused Note: In HPOG-Impact data collection, education/credential items text changed to measure completion since random assignment	Exploratory Outcome	15-month follow-up survey Q25a and Q25b
Educational Progress	Perception of progress towards long-range educational goals	Indicator variable that takes on a value of 1 if the person strongly or somewhat agrees with the statement "I am making progress towards my long-range educational goals", and a value of 0 if he/she strongly or somewhat disagrees.	Exploratory Outcome	15-month follow-up survey Q35
Employment	Proportion of quarters employed during the five quarters after the quarter containing random assignment	1. Create binary variables that indicate employment during each of the quarters 2. Average the quarters	Exploratory Outcome	NDNH

Domain	Variable Description	Operationalization/Details	Outcome Designation	Data Source(s) Items
Employment in Healthcare	Currently employed in a healthcare job	Binary variable that takes on a value of 1 if the individual is currently employed AND works for a healthcare employer AND is employed in a healthcare job Missing if don't know/refused 0 otherwise	Exploratory Outcome	15-month follow-up survey Q36, Q36d
Self-Efficacy & Motivation	General Self-Efficacy Scale (GSE) based on Schwarzer & Jerusalem (1995)	1. Assign the following values to the response to each item: 1 "Not at all true" 2 "Somewhat true" 3 "Mostly true" 4 "Entirely true" (Leave don't know & refused as missing) 2. Average the observed items. 3. Replace the average with missing if at fewer than 6 of the 9 items are observed	Exploratory Outcome	15-month follow-up survey Q47

Exhibit A.9: Program Components

Domain	Variable Description	Operationalization/Details	Data Source(s) Items
Presence of Career Pathways Principles	Extent to which available offerings and program content is based on principles of the career pathways framework	Continuous variable ranging from 0-8 that counts the number of elements of the career pathways framework that are being implemented by the program. Assign one point for each of the following elements that are available or readily available: * Opportunities that emphasize career pathways * Opportunities that target individuals with significant skill, education, and work experience deficits * Curricula that accommodate multiple learning modes and capabilities * Opportunities designed to accommodate non-traditional student populations * Opportunities to orient and acclimate non-traditional student populations to health professions Assign one point for each of these elements that are offered: * training options that provide credentials that are “stackable” with other available training * a set of training options that support multiple career pathways * health or vocational education/training activities designed (or redesigned/compressed) for accelerated completion	Grantee Survey 4.1 8.7 8.10
Case Management	Average caseload for FTE (estimated full time equivalent) case managers	Continuous variable constructed by taking average of: -Average Estimated Caseload for full-time case managers -Double the Average Estimated Caseload for part-time case managers	Grantee Survey 9.3
Case Management	Number of services that case managers and counselors deliver that meet the needs of participants	Continuous variable that counts the number of services delivered by case managers that the program agrees or strongly agrees meet their participants’ needs: -Participant monitoring -Academic counseling -Career counseling -Counseling to identify personal and supportive service needs -Financial counseling -Job search/placement assistance -Job retention services	Grantee Survey 9.2 9.6

Domain	Variable Description	Operationalization/Details	Data Source(s) Items
Comprehensive Services	Access to social and other services: social and other services delivered that meet participants' needs	Continuous variable that assigns two points for each service directly delivered by a program and one point for each service for which the program delivers a referral (three points possible per service). A point is deducted from this sum for each of these services that the program does not agree or strongly agree meets their participants' needs. These services include: -Mentoring Activities -Cultural Programming -Driver's license assistance -Food assistance (other than SNAP) -Addiction or substance abuse services -Family preservation services -Family engagement services -Legal assistance -Primary or Medical Care -Short-term/temporary housing -Other housing assistance	Grantee Survey 9.8 9.10 9.11 9.12

Domain	Variable Description	Operationalization/Details	Data Source(s) Items
Comprehensive Services	Access to and delivery of tuition and other financial services: tuition coverage plus financial services offered that meet participant needs	Continuous variable from 0-2 based on two components: 1) Percentage of tuition covered by HPOG and 2) Financial services delivered that meet participant needs. Programs that cover 100% tuition and offer all possible financial services receive the maximum of 2. Sums two constructs assigned as follows: -Tuition coverage: Up to one point for programs where tuition is fully covered. Quarter, 0.5, or 0.75 points assigned for partial tuition coverage. -Financial Services: Three points for each service provided to all participants, 2 for those provided upon request or with a course, 1 for those provided with no other specifics. A point is deducted from this sum for each of these services that the program does not agree or strongly agree meets their participants' needs. This sum is divided by 15 (the maximum possible) to create a 0-1 score. Possible financial services include: -Book costs -Licensing and certification fees -Exam/exam preparation fees -Work/training uniforms, supplies, tools -Computer/technology equipment	Grantee Survey 9.14 9.17 9.18 9.20
Comprehensive Services	Access to childcare and transportation: accessibility via public transportation plus childcare and transportation services offered that meet participant needs	Continuous variable based on two components: 1) Percentage of catchment area with access to public transportation and 2) Childcare and transportation services delivered that meet participant needs Sums two constructs assigned as follows: -Two points assigned if entire catchment area has access to public transportation, one point if 75% has access, zero otherwise -Two points for each support directly delivered by a program and one point for each service for which the program delivers a referral. A point is deducted from this sum for each of these supports that the program does not agree or strongly agree meets their participants' needs. These supports include: -Child care assistance -Transportation assistance	Grantee Survey 2.3a 9.11 9.12

Domain	Variable Description	Operationalization/Details	Data Source(s) Items
Comprehensive Services	Location of services: number of services co-located with the training site	Continuous variable that counts the number of services that are physically co-located with the healthcare training most or all of the time. Services that may be co-located: -Academic advising/counseling -Financial aid advising/counseling -Advising/counseling about support services -Career advising/counseling -Job placement services -Basic skills instruction, GED preparation, ESL, or other training activities	Grantee Survey 3.7
Employment Supports	Number of employment supports that are offered that meet participants' needs	Continuous variable that assigns two points for each employment support service directly delivered by a program and one point for each service for which the program delivers a referral (three points possible per service). A point is deducted from this sum for each of these services that the program does not agree or strongly agree meets their participants' needs. These services include: -Job-readiness workshops -Job search skills workshops -Identifying job openings for program graduates -Meeting with employers to identify job openings for graduates -One-on-one job search assistance -Advising on career and job choices -Operating or referrals to job fairs -Providing participants with job listings -Job screening -Post-placement services	Grantee Survey 9.24 9.26
Behavioral Incentives	Non-cash incentives: whether the program provides offer non-cash incentives to participants for achieving program milestone	Binary variable 1 if non-cash incentives are perceived to be effective or very effective in encouraging participants to achieve the desired program milestones 0 if perception of non-cash incentives effectiveness are neutral, not effective, or if they are not offered	Grantee Survey 8.21
Peer Support	Offer of facilitated peer support	Continuous variable that assigns two points if peer support activities are directly delivered by the program and one point if the program delivers a referral. A point is deducted if the program does not agree or strongly agree that they are able to meet their participants' needs for peer support.	Grantee Survey 9.8 9.10

Domain	Variable Description	Operationalization/Details	Data Source(s) Items
Emergency Assistance	Access to emergency funds to meet needs stemming from imminent eviction from housing, utility shutoff, vehicle repair needs, etc.	Continuous variable that counts the number of emergency services delivered that the program agrees or strongly agrees meet their participants' needs: -Car repair costs -Car insurance costs -Utilities (e.g., heating, electricity, water bills) -Food assistance (non-SNAP) -Security deposit -Rent -Housing Program fees	Grantee Survey 9.19 9.20

Exhibit A.10: Implementation Features

Domain	Variable Description	Operationalization/Details	Data Source(s) Items
Management/Staff Focus	Extent to which program is employment or education focused	Two continuous measures created to summarize: - The percentage of management/staff that indicate employment is the primary goal of the program (response categories 1, 2 and 3) - The percentage of management/staff that indicate education is the primary goal of the program (response categories 5, 6, 7) The percentage of management and staff members that indicate both education and employment as the goal of the program is omitted	Staff & Management Survey 27
Staff Experience	Percentage of management/staff at the division level with at least five years of experience	Percentage of respondents within the division that indicate they have greater than five years experiences.	Staff & Management Survey 2a
Staff Discretion/Autonomy	Staff perception of autonomy, including authority to carry out responsibility, ability to try different techniques, trust in staff professional judgment and not too many rules.	Individual staff/manager index created by averaging Likert scales measuring on agreement with the following statements: -Staff in your program are given broad authority in carrying out their responsibilities -Staff in your program can try out different techniques to improve their effectiveness -Staff members are given too many rules in your program (REVERSE CODED) -Management/You fully trusts/trust the professional judgments of staff in your program Average the responses at the division level	Staff & Management Survey D80 D81 D82 D83

Exhibit A.11: Individual-Level Measures of Program Participation at 15-Month Follow-Up

Domain	Variable Description	Operationalization/Details	Data source(s)
Education	Obtained a professional, state or industry certificate, license or credential since random assignment	Binary variable with value 1 if individual reports: * Obtained a professional, state or industry certificate, license or credential Value 0 if individual reports either of the following: * No formal training * Some coursework or training but no professional, state or industry certificate, license or credential Missing if don't know/refused Note: In HPOG-Impact data collection, education/credential items text changed to measure completion since random assignment	15-month follow-up survey Q25b
Education	Completed a degree (AA, BA or higher) or obtained a credential (professional, state or industry certificate, license or credential) since random assignment	Binary variable with value 1 if one or both of the following is true: * Completed an Associate's degree since RAD * Completed a Bachelor's degree or above since RAD Value 0 otherwise Missing if don't know/refused Note: In HPOG-Impact data collection, education/credential items text changed to measure completion since random assignment	15-month follow-up survey Q25a
Emergency Assistance	Received emergency assistance	Binary variable with value 1 if individual reports receiving: * Emergency assistance, or funds to cover the costs of unexpected personal crisis, such as utility shut off or car repair Value 0 otherwise Missing if don't know/refused	15-month follow-up survey Q28j
Non-Cash Incentive	Received a non-cash incentive	Binary variable with value 1 if individual reports receiving: * Incentives, for example, a gift card for completing a course Value 0 otherwise Missing if don't know/refused	15-month follow-up survey Q28k
Peer Support	Participated in peer-support	Binary variable with value 1 if individual reports participating in: * Peer support groups Value 0 otherwise Missing if don't know/refused	15-month follow-up survey Q28i

Domain	Variable Description	Operationalization/Details	Data source(s)
Personal Counseling	Used personal counseling services	Binary variable with value 1 if individual reports receiving: * Personal counseling Value 0 otherwise Missing if don't know/refused	15-month follow-up survey Q28h
Academic Assistance	Used academic assistance services (tutoring)	Binary variable with value 1 if individual reports receiving: * Tutoring in subjects where you needed extra help Value 0 otherwise Missing if don't know/refused	15-month follow-up survey Q28c

Exhibit A.12: Metropolitan Statistical Areas (MSAs) Served by Each HPOG Program

Site	MSAs Served
Buffalo and Erie County WDC (Buffalo, NY)	Buffalo-Niagara Falls, NY
NH OMHRA (Concord, NH)	Other New Hampshire nonmetropolitan area (18.0%) Manchester, NH (28.5%) Nashua, NH-MA NECTA Division (20.7%) Western New Hampshire nonmetropolitan area (14.2%) Rochester-Dover, NH-ME (18.6%)
Research Foundation of CUNY (Bronx, NY)	New York-White Plains-Wayne, NY-NJ Metropolitan Division
Schenectady County Community College (Schenectady, NY)	Albany-Schenectady-Troy, NY
Suffolk County Department of Labor (Hauppauge, NY)	Nassau-Suffolk, NY Metropolitan Division
The Workplace, Inc. (Bridgeport, CT)	Bridgeport-Stamford-Norwalk, CT
Alamo Community College (San Antonio, TX)	San Antonio-New Braunfels, TX
KS Department of Commerce (Topeka, KS)	Topeka, KS
KS LWIB I (Salina, Kansas)	Kansas nonmetropolitan area (94.5%) Wichita, KS (5.5%)
KS LWIB II (Heartland Works, Inc.: based in Topeka, Lawrence, Manhattan, and Junction City, KS)	Topeka, KS (36.3%) Manhattan, KS (37.4%) Lawrence, KS (26.3%)
KS LWIB III (Workforce Partnership: Kansas City, Overland Park and Leavenworth, KS)	Kansas City, MO-KS
KS LWIB IV (Workforce Alliance of South Central Kansas, located in Wichita)	Wichita, KS
KS LWIB V (KANSASWORKS Southeast Region Kansas Health Profession Opportunity Project (KHPOP): Located in Independence (city) and Montgomery (county))	Kansas nonmetropolitan area
Full Employment Council (Kansas City, MO)	Kansas City, MO-KS
Gateway Community & Tech College (Florence, KY)	Cincinnati-Middletown, OH-KY-IN
WIB SDA83 Inc. (Monroe, LA)	Monroe, LA (53.4%) Winnsboro nonmetropolitan area (46.6%)
Will County WIB (Joliet, IL) Central States SER College of Lake County Jewish Vocational Service Joliet Junior College	Chicago-Joliet-Naperville, IL Metropolitan Division
Edmonds Community College (Lynnwood, WA)	Seattle-Bellevue-Everett, WA Metropolitan Division
Bergen Community College (Paramus, NJ)	New York-White Plains-Wayne, NY-NJ Metropolitan Division
Brookdale Community College	Edison-New Brunswick, NJ Metropolitan Division
Essex Community College	Newark-Union, NJ-PA Metropolitan Division
Hudson Community College	New York-White Plains-Wayne, NY-NJ Metropolitan Division
Meridian Health	Newark-Union, NJ-PA Metropolitan Division
Middlesex Community College	Edison-New Brunswick, NJ Metropolitan Division
Community College of Morris	Newark-Union, NJ-PA Metropolitan Division
Passaic Community College	New York-White Plains-Wayne, NY-NJ Metropolitan Division
Sussex Community College	Newark-Union, NJ-PA Metropolitan Division
Union Community College	Newark-Union, NJ-PA Metropolitan Division
Warren County Community College	Allentown-Bethlehem-Easton, PA-NJ

Site	MSAs Served
Central Susquehanna Intermediate Unit (Lewisburg PA)	East Central Pennsylvania nonmetropolitan area (29.5%) Northeastern PA nonmetropolitan area (11.6%) West Central PA nonmetropolitan area (9.5%) State College, PA (8.4%) Williamsport, PA (41.1%)
Pensacola State College (Pensacola, FL)	Pensacola-Ferry Pass-Brent, FL
South Carolina Department of Social Services (Columbia, SC)	Columbia, SC (22.3%) Florence, SC (23.3%) Greenville-Mauldin-Easley, SC (22.5%) Low Country South Carolina nonmetropolitan area (32.0%)
Central Community College (Grand Island, NE)	Central Nebraska nonmetropolitan area (71.9%) Northeastern Nebraska nonmetropolitan area (13.6%) Western Nebraska nonmetropolitan area (14.5%)
Eastern Gateway Community College (Steubenville, OH)	Steubenville-Weirton, OH-WV (26.2%) Youngstown-Warren-Boardman, OH-PA (37.2%) Other Ohio nonmetropolitan area (36.5%)
Milwaukee Area WIB (Milwaukee, WI)	Milwaukee-Waukesha-West Allis, WI

The following are the sources we use to construct the grantee-level measures of local labor market information (including health sector employment and wage conditions):

- BLS Occupational and Employment Statistics (OES) data provide employment and wage information by detailed occupation for Metropolitan Statistical Areas (MSAs), Non-MSAs, and States (see: http://www.bls.gov/oes/#data)
- BLS Unemployment Statistics for States and MSAs (see: http://www.bls.gov/lau/#tables)

If a given HPOG location serves individuals from only one MSA or non-MSA, then it is straightforward to link BLS data to that HPOG location. As depicted by Exhibit A.8, there are some HPOG programs that serve multiple MSAs. The Area Served Column indicates how study participants from a given site are distributed across MSAs. We link HPOG sites that cover multiple MSAs to BLS data using a weighted combination of BLS data from all MSAs served by the HPOG program, where the weights are the percent of HPOG study participants in each MSA within an HPOG program. Note that for BLS Unemployment statistics, we will use the state average unemployment rate for non-MSA areas (because the BLS does not report unemployment rates for non-MSAs).

Appendix B. Plan for Calculating Attrition

The following details the steps in our plan for calculating individual-level attrition:

- Define the base number of treatment individuals as Base.T = number of randomly assigned treatment individuals
- Define the base number of control individuals as Base.C = number of randomly assigned control individuals
- Define the number of treatment individuals included in the final analytic sample for a given contrast as Assessed.T = number of treatment individuals included in the final analytic sample (e.g., those with non-missing outcome data)
- Define the number of control individuals included in the final analytic sample for a given contrast as Assessed.C = number of control individuals included in the final analytic sample
- Define the overall attrition rate as Attrition.O = 1 –[(Assessed.T+ Assessed.C)/(Base.T+ Base.C)]
- Define the attrition rate in treatment group as Attrition.T = 1 – (Assessed.T/ Base.T)
- Define the attrition rate in control group as Attrition.C = 1 – (Assessed.C/ Base.C)
- Define differential attrition as Attrition.D = |Attrition.T- Attrition.C|

Exhibit B.1 suggests a table shell for reporting attrition rates.

Exhibit B.1: Overall and Differential Attrition Rates

A	B	C	D	E	F	G	H
Contrast ID #	Contrast Name (optional)	Treatment Group # Individuals Randomized	Control Group # Individuals Randomized	Treatment Group # Individuals in final Analysis Sample	Control Group # Individuals in final Analysis Sample	Overall Attrition Rate for the entire sample	Differential Attrition Rate (difference in rates of attrition for treatment and control groups)

Appendix C. Sample Sizes and Missing Data Rates for Description of Study Sample

Exhibit C.1: Sample Sizes and Missing Observations for Demographic Characteristics of Sample at Baseline

Characteristic	Entire Sample	Treatment Group	Control Group
Sex			
N	13,570	8,599	4,971
Missing Rate (%)	0.0	0.0	0.1
Marital Status			
N	12,955	8,256	4,699
Missing Rate (%)	4.6	4.0	5.5
Dependent Children			
N	13,006	8,321	4,685
Missing Rate (%)	4.2	3.3	5.8
Race and ethnicity			
N	13,313	8,496	4,817
Missing Rate (%)	1.9	1.2	3.2
Age			
N	13,545	8,583	4,962
Missing Rate (%)	0.2	0.2	0.2
Born Outside the U.S.			
N	13,313	8,496	4,817
Missing Rate (%)	1.9	1.2	3.2

Exhibit C.2: Sample Sizes and Missing Observations for Educational Background of Sample at Intake

Characteristic	Entire Sample	Treatment Group	Control Group
Educational Attainment			
N	13,280	8,483	4,797
Missing Rate (%)	2.2	1.4	3.6
Credential/Degree completion			
Post-secondary degree/certificate			
N	11,631	7,406	4,225
Missing Rate (%)	14.3	13.9	15.1
Occupational Skills License or Certification			
N	12,350	7,874	4,476
Missing Rate (%)	9.0	8.5	10.0
Literacy at the 8th Grade Level or Higher			
N	8,893	6,137	2,756
Missing Rate (%)	16.2	13.8	21.3
Numeracy at the 8th Grade Level or Higher			
N	8,586	5,927	2,659
Missing Rate (%)	19.1	16.8	24.1
Previous Preparation Classes			
Adult Basic Education			
N	13,053	8,255	4,798
Missing Rate (%)	3.8	4.0	3.5
English as a Second Language			
N	13,068	8,268	4,800
Missing Rate (%)	3.7	3.9	3.5
Classes in how to succeed in school			
N	13,049	8,258	4,791
Missing Rate (%)	3.9	4.0	3.7
Classes in how to succeed in work			
N	13,012	8,229	4,783
Missing Rate (%)	4.1	4.3	3.8
Vocational, Technical or Trade School			
N	13,187	8,349	4,838
Missing Rate (%)	2.9	2.9	2.7
Level Expected to Complete			
N	13,166	8,368	4,798
Missing Rate (%)	3.0	2.7	3.5

Exhibit C.3: Sample Sizes and Missing Observations for Economic Status of Sample at Baseline

Characteristic	Entire Sample	Treatment Group	Control Group
Household Income			
N	11,831	7,554	4,277
Missing Rate (%)	12.8	12.2	14.0
Individual Income			
N	9,877	6,695	3,182
Missing Rate (%)	7.0	6.0	9.2
Public Assistance Use			
WIC/SNAP (% Receiving)			
N	13,189	8,417	4,772
Missing Rate (%)	2.8	2.1	4.1
Welfare (% Receiving)			
N	13,049	8,334	4,715
Missing Rate (%)	3.9	3.1	5.2

Exhibit C.4: Sample Sizes and Missing Observations for Employment and Expected Time Use of Sample at Baseline

Characteristic	Entire Sample	Treatment Group	Control Group
Current Working Hours			
N	12,643	8,070	4,573
Missing Rate (%)	6.9	6.2	8.1
Expected Working Hours			
N	12,539	7,960	4,579
Missing Rate (%)	7.6	7.5	7.9
Expect to Participate in HPOG			
N	13,006	8,257	4,749
Missing Rate (%)	4.2	4.0	4.5

Exhibit C.5: Sample Sizes and Missing Observations for Life Challenges of Sample at Baseline

Characteristic	Entire Sample	Treatment Group	Control Group
Limited English Proficiency (%)			
N	13,057	8,344	4,713
Missing Rate (%)	3.8	3.0	5.2
Barriers to school/work (% at fairly or very often)			
Child care arrangements			
N	13,154	8,336	4,818
Missing rate (%)	3.1	3.1	3.1
Transportation			
N	13,188	8,353	4,835
Missing rate (%)	2.9	2.9	2.8
An illness or health condition			
N	13,370	8,481	4,889
Missing rate (%)	1.5	1.4	1.7
Alcohol or drug use			
N	13,447	8,541	4,906
Missing rate (%)	0.9	0.7	1.4

Appendix D. Process for Selecting Covariates for Section 6.2 Model

To select the candidate measures to be included in the model, we will use the following forward stepwise procedure:

Step 1: Estimate Equation (6-14) including individual-level baseline covariates and the indicator for offering facilitated peer support. These covariates are the starting set of standard covariates. Compute the estimated bias using Equation (6-16).

Step 2: Estimate Equation (6-14) while including all of the standard covariates plus one candidate measure (see Section 4.1 for list of candidate measures). Compute the estimated bias using Equation (6-16). Repeat this step for each of the candidate measures listed in Section 4.1.

Step 3: Rank all of the estimated biases computed in Step 2.

- If the lowest estimated bias produced in Step 2 is higher than the estimated bias produced in Step 1, stop and do not include any candidate measures in the model used to report findings.
- If the lowest estimated bias produced in Step 2 is lower than the estimated bias produced in Step 1, add the covariate associated with that estimated bias to the list of standard (automatically included) covariates. Go on to Step 4.

Step 4: Estimate Equation (6-14) while including all standard covariates plus one of the remaining candidate measures. Compute the estimated bias. Repeat this step for each of the remaining candidate measures.

Step 5: Rank all of the estimated biases produced in Step 4.

- If the lowest estimated bias produced in Step 4 is higher than the lowest estimated bias produced up to this point, stop and do not include any additional candidate measures in the model used to report findings.
- If the lowest estimated bias produced in Step 4 is lower than the lowest estimated bias produced up to this point, add the covariate associated with that estimated bias to the list of standard (automatically included) covariates. Go on to Step 6.

Step 6: Let "D" represent the number of division-level factors included in the mode; let "P" represent the number of program-level factors included in the model; and let "L" represent the number of local context-level measures included in the model. We impose the following constraint on the maximum number of division- and program- level covariates to be included in the model: $14 \geq D+(14/7)P+(14/5)L$.[33] Repeat Steps 4 and 5 until no additional covariate is added in any one cycle or until the maximum number of covariates is reached. This yields the version of Equation (6-14)—and hence the Section 6.2 model overall—from which reported findings will be derived.

[33] This ensures that we include no more than one division-level predictor for every 5 divisions included in the Equation (6-14) model; no more than one program-level predictor for every 5 programs included in the model; and no more than one local context-level predictor for every 5 local context areas included in the model. These degrees of freedom are based on the modeling sample and are slightly more restrictive than those reflected in Exhibit 6.2, which applies a similar equation to the entire sample.

Appendix E. Source of Omitted Variable Bias in Non-experimental Estimates

This appendix formalizes how omitted-variable bias arises when producing non-experimental estimates in the basic multi-level model described in Section 6.2 of the main text. Consider first the following simplified two-level model that (i) aggregates divisions within programs and (ii) suppresses notation for individual characteristics (IC_c), participant composition variables (PC_d), and local context variables (LC_q):

Level One: Individuals

$$Y_{ki} = \alpha_k + \beta_k T_{ki} + \varepsilon_{ki} \quad \text{(eq. E-1)}$$

where:

- Y_{ki} = the outcome measure for individual *i* (from division *j*, not shown) in program *k*;
- T_{ki} = the standard HPOG program treatment indicator (1 for those individuals assigned to the standard HPOG treatment; 0 for the control group individuals);
- α_k = the control group mean outcome (counterfactual) for program *k*;
- β_k = the conditional impact of being offered the standard HPOG program *k*;
- ε_{ki} = a random component of the outcome for each individual.

Level Three: Programs

$$\beta_k = \beta_0 + \sum_m \pi_m P_{mk} + \sum_g \varphi_g \bar{I}_{gk} + \mu_k \quad \text{(eq. E-2)}$$

where:

- P_{mk} = program component m for program k (grand mean centered), m = 1, . . ., M including the experimentally varied enhancement components P_{Sk}, P_{Ak}, and P_{Ik};
- $\bar{I}_{gk}$ = implementation feature g averaged across divisions within program k (grand mean centered);
- β_0 = the grand mean impact;
- π_m = the influence of program component m on impact magnitude, m = 1, . . ., M;
- φ_g = the influence of implementation feature g on impact magnitude, g = 1, . . ., G; and
- μ_k = a random component of intervention impacts for each program.

$$\alpha_k = \alpha_0 + v_k \quad \text{(eq. E-3)}$$

where:

α_0 = the grand mean control group outcome;

v_k = a random component of control group mean outcome for each program.

Estimating this model will produce estimates $\hat{\pi}_m^N$ of the relationship between program component *m* and treatment impact magnitude for each of the *m = 1,..., M* program features and estimates $\hat{\varphi}_g^N$ of the relationship between implementation feature *g* and treatment impact magnitude for each of the *g = 1,..., G* implementation features. The *N* superscript denotes that these estimates are non-experimental.

Omitted variable bias can arise when estimating $\hat{\pi}_1^N, \ldots, \hat{\pi}_M^N$ and $\hat{\varphi}_1^N, \ldots, \hat{\varphi}_G^N$ if:

1. There is an omitted program-level factor F_k(which could be the aggregation to the program level of an omitted division-level factor) that influences treatment impact magnitudes;
2. F_k is not included among the program components $P_{1k}, P_{2k}, \ldots, P_{Mk}$ or the implementation features $\bar{I}_{1k}, \bar{I}_{2k}, \ldots, \bar{I}_{Gk}$ (or the suppressed PC_d or LC_q variables) in the model; and
3. F_k correlates with one or more of the included program components P_{mk} or implementation features $\bar{I}_{gk}$.

Collectively, these characteristics of F_k make it an unmeasured confounder in the analysis. Confining attention to program components (the remainder of the argument holds for implementation features as well, which are now suppressed from the notation), omission of such an F_k factor leads to biased estimates of π_m^N whenever $Cov(P_{mk}, F_k) \neq 0$. Assuming that the influence of F_k is linear and separable from the influence of the other factors (i.e., there is no interaction between it and the included intervention features when determining impact magnitude), the true version of Equation (E-2) (with the φ_g terms suppressed) is

$$\beta_k = \beta_0 + \sum_m \pi_m P_{mk} + \lambda F_k + \mu_k, \qquad \text{(eq. E-2')}$$

where:

λ = the amount by which a one-unit change in F (an omitted confounder) alters impact magnitude (this is "lambda").

To see how the bias arises, plug Equation (E-2') and Equation (E-3) into Equation (E-1) to get:

$$Y_{ki} = \alpha_0 + \beta_0 T_{ki} + \sum_m \pi_m P_{mk} T_{ki} + \lambda F_k T_{ki} + \{ \mu_k T_{ki} + \varepsilon_{ki} + v_k \} \qquad \text{(eq. E-4)}$$

When this equation is estimated with maximum likelihood methods (Bryk and Raudenbush, 1992)[34] with F_j omitted, the probability limit of the resulting estimators of the π_m coefficients, call them $\hat{\pi}_m^N$, is:[35]

$$plim(\hat{\pi}_m^N) = \pi_m + \lambda \frac{\text{Cov}(P_m, F)}{\text{Var}(P_m)}, m = 1, \ldots, M, \qquad \text{(eq. E-5)}$$

[34] More specifically, when maximum likelihood methods are applied to the multi-level system of Equations (E-1), (E-2'), and (E-3) from which Equation (E-4) is derived.

[35] Once again, the *N* superscript here denotes that these estimates attribute impact to different causal factors nonexperimentally.

from the usual econometric formula for omitted-variable bias at the limit as sample size goes infinite.[36] The bias of $\hat{\pi}_m^N$ from this probability limit can be expressed by the following equation:

$$bias(\hat{\pi}_m^N) = plim(\hat{\pi}_m^N) - \pi_m = \lambda \frac{\text{Cov}(P_m, F)}{\text{Var}(P_m)}. \quad \text{(eq. E-6)}$$

If $Cov(P_m, F) \neq 0$, then $\hat{\pi}_1^N, \ldots, \hat{\pi}_M^N$ from the misspecified model in Equations (E-1), (E-2), and (E-3) are biased estimates for $\pi_1, \ldots, \pi_M$ even in very large samples, with bias that approaches $\lambda \frac{\text{Cov}(P_m, F)}{\text{Var}(P_m)}$ as sample size goes infinite for $m = 1, \ldots, M$.

[36] See for example Wooldridge (2002), pp. 61-62. As Wooldridge notes, this expression for bias makes common simplifying assumptions.