HERKIMER, MADISON, AND ONEIDA COUNTIES WORKFORCE DEVELOPMENT BOARD

BRIDGE TO EMPLOYMENT AND ACADEMIC MARKETPLACE

WORKFORCE INNOVATION FUND

FINAL EVALUATION REPORT: IMPACT STUDY

SEPTEMBER 2019

Submitted by: The Policy & Research Group www.policyandresearch.com

> 8434 Oak St. New Orleans, LA 70118

1411 4th Ave., Suite 1000 Seattle, WA 98101



This report presents methods and analyses from PRG's impact evaluation of the Bridge to Employment and Academic Marketplace. Contents are excerpted from the full evaluation report that was submitted to the U.S. Department of Labor's Workforce Innovation Fund on September 20, 2019.

TABLE OF CONTENTS

BEAM Program Design	1
Impact Evaluation	3
Research Questions	3
Design and Analytic Procedures	4
Findings	9
Discussion	
Employment and Wages	20
Conclusions and Study Limitations	20
Appendix A: Impact Evaluation Methods	22
Purpose and Background	22
Research Questions	23
Methods	
Analytic Samples	
Benchmark Analytic Results	45
Appendix B: Impact Evaluation Sensitivity Studies	51
Description of Sensitivity Studies	51
Sensitivity Study Analytic Results	

BEAM PROGRAM DESIGN

Bridge to Employment and Academic Marketplace (BEAM) was an initiative funded by a United States Department of Labor (USDOL) Workforce Innovation Fund (WIF) Round 2 grant, which was implemented November 17, 2015 through September 30, 2018.¹ BEAM was created to bridge the gap between workforce and education systems by providing in-depth services to individuals disconnected from work or education. The services included community outreach and recruitment, providing academic and career-related resources, and engaging in customized one-on-one coaching for enrolled individuals. These services were intended to help economically disadvantaged adults return to and complete postsecondary education and/or training. The program had three broad goals:

- 1. To minimize the time that participants spend disconnected from school and connect college dropouts, or stop-outs, with programs that meet both their career goals and the region's job needs;
- 2. To build a system where the workforce and education systems collaborate to recruit for postsecondary enrollment that is ultimately essential for good-paying jobs; and
- 3. To create a replicable service model at the program level that others can learn from and adapt.

BEAM was a partnership across a consortium of four workforce development boards in New York: lead grantee Herkimer, Madison, and Oneida Counties Workforce Development Board (HMO WDB) and partner grantees Broome-Tioga Workforce Development Board (Broome-Tioga WDB); Chenango, Delaware, and Otsego Counties Workforce Development Board (CDO WDB); and Tompkins County Workforce Development Board (Tompkins WDB).

Through this partnership, the BEAM program aimed to enroll college dropouts, or stop-outs, in postsecondary training and/or education that leads to high-growth jobs. The BEAM program used Outreach Coordinators (OCs) as a one-stop resource for BEAM participants, providing customized one-on-one coaching, academic, and employment support services throughout a participant's experience in the program.

Individuals were eligible to participate in BEAM if they:

- 1. were 19 years or older,
- 2. resided in one of nine participating counties,
- 3. had previously attempted postsecondary education but dropped out before receiving a degree or certificate,
- 4. gave consent to participate in the Impact Study,
- 5. were determined by OCs to be a good fit for the study (i.e., likely to return to an American Job Center (AJC) to receive services),
- 6. enrolled to receive services at one of the participating AJC during the study enrollment period,
- 7. were not a veteran of the armed forces,² and
- 8. were not currently enrolled in a postsecondary education program.

¹ Round 2 USDOL WIF grants were implemented in three phases: Phase I focused on grant preparation (October 1, 2014 – September 30, 2015), Phase II was grant implementation (October 1, 2015 – September 30, 2018), and Phase III focused on data collection and evaluation (October 1, 2018 – September 30, 2019). Phase I came to a close for BEAM once the program's final work plan was approved by USDOL on November 17, 2015.

² As outlined in the WIF grant funding requirements, veterans of the armed forces were excluded from participation in the randomized controlled trial, but veterans were eligible to receive BEAM services.

The following are specific services that were available through BEAM to participants³:

- One-on-one OC navigation, advocacy, and customized assistance through entire process from re-enrollment to employment. This included mentoring and encouragement.
- Individual Academic Plan development and revisions and individual academic goal planning/career counseling.
- Direct postsecondary education financial aid support from OCs (e.g., counseling, navigation of institutions' financial aid offices, one-on-one assistance with completing applications, navigation of default loan rehabilitation).
- Regular contact with OC (or partner staff), ranging from daily to monthly.
- Referrals to community resources.

Those services were in addition to ones available to AJC customers and the general public:

- Job matching and career assessment tools provided at AJCs.
- Income, health, and family support programs such as TANF, SNAP, Medicaid, and childcare subsidies.
- Remedial education support through Educational Opportunity Center (EOC), Board of Cooperative Educational Services (BOCES), and other community-based organizations.
- Financial aid support and assistance through institutions' financial aid offices or communitybased organizations.
- TABE testing and services.
- Basic skills development sessions provided by community partners.
- Academic tutoring services, workshops, or information sessions provided at community colleges or community sites.
- Resume building assistance through community partner organization or community colleges' career services offices.
- Internship and apprenticeship opportunities through community colleges' career services offices and workforce system.
- On the Job Training placement through workforce system.

³ As discussed later in this section and in the Impact Evaluation section, these BEAM-specific services were provided to the Guided Career Pipeline. Another group, the Career Center Services did not receive these services.

IMPACT EVALUATION

The purpose of PRG's Impact Evaluation was to assess whether BEAM improved customer educational and employment outcomes. The study employed a randomized controlled trial (RCT) design to answer two primary and four exploratory research questions that were concerned with BEAM's effect on educational and employment outcomes. As detailed in the Evaluation Design Report (EDR), the aim was to isolate the causal impact that BEAM had on an individual's likelihood of returning to and persisting through a postsecondary education program, obtaining a degree or certificate, entering employment, and its effect on wages.

The Evaluation Team's approach was to assess the impact of BEAM by means of a RCT, where eligible individuals were randomly assigned to one of two conditions: (1) the offer to participate in the innovative set of BEAM services, referred to as the Guided Career Pipeline (GCP) or treatment condition, or (2) the offer to participate in American Job Center (AJC) services as usual, referred to as the Career Center Services (CCS) or control condition. The Impact Study contrasted the employment and educational outcomes of individuals randomly assigned to the treatment condition (GCP) with those of individuals assigned to receive services that they would typically be offered by the participating AJCs (control condition).

BEAM was a program model that employed a core network of staff to deliver individualized assistance to reenroll disconnected youth into education and employment opportunities that went above and beyond what was traditionally provided at AJCs prior to its inception. Individuals enrolled in GCP (treatment) were offered individualized support from OCs but could have also benefited from traditional services and funding under other federal programs available to CCS participants. The control group (CCS) was only offered the traditional suite of services. The Impact Study aimed to assess the educational and employment impact of being offered GCP (treatment) relative to being offered CCS (control), which is a business-as-usual contrast. The impact of BEAM's GCP program on educational and employment outcomes was estimated within the intent-to-treat (ITT) framework, which means that the analysis included all of the participants initially enrolled into the study, regardless of their actual exposure to the intervention they were assigned to.

Educational outcomes were measured using data from the National Student Clearinghouse (NSC), and employment outcomes were measured using data from the New York State Department of Labor (NYSDOL). The impact of the GCP program is estimated with a regression equation that models outcomes as a function of treatment status, the baseline measure of the outcome variable (or a proxy), and other covariates. An ordinary least squares (OLS) model was used to estimate all outcomes.

RESEARCH QUESTIONS

By design, the Impact Evaluation was developed to answer two primary and four exploratory research questions regarding BEAM's effect on outcomes identified by the program's theory of change. The primary study focuses on the program's short-term educational outcomes (postsecondary enrollment and persistence), whereas the exploratory study focuses on the program's indirect educational and economic outcomes (graduation from postsecondary program and improved employment outcomes).⁴

⁴ Additional explanation of the distinction between the primary and exploratory outcomes of the Impact Study is provided in Appendix A.

PRIMARY STUDY

- 1. Are treatment group members who are offered Guided Career Pipeline (GCP) more likely to enroll in a postsecondary education program within nine months of enrolling in BEAM than equivalent participants who were offered Career Center Services (CCS)?
- 2. Do treatment group members who are offered GCP remain enrolled in a postsecondary education program for more consecutive semesters than equivalent participants who were offered CCS?

EXPLORATORY STUDY

- 3. Are treatment group members who are offered GCP more likely to complete their postsecondary program certificate or degree by the end of the study period than equivalent participants who were offered CCS?
- 4. Are treatment group members who are offered GCP more likely to enter employment within three quarters of enrolling in BEAM than equivalent participants who were offered CCS?
- 5. Do treatment group members who are offered GCP receive higher wages at the end of the 3rd quarter post-enrollment than equivalent participants who were offered CCS?
- 6. Do treatment group members who are offered GCP remain employed for more quarters on average than equivalent participants who were offered CCS?

As described in the Program Design section of this report, BEAM aimed to reconnect unemployed, outof-school youth and adults with postsecondary education and training opportunities that would lead to industry-recognized credentials and employment in high-growth and living-wage careers. The primary study examines the impact of BEAM's GCP program on the likelihood that participants will enroll in a postsecondary education training program and will remain enrolled (persist) in their chosen programs for a longer period of time. The exploratory study aims to examine the impact of the GCP program on the exploratory outcomes of degree attainment, increased employment, and higher wages. These outcomes are considered secondary because the program's theory of change does not explicitly hypothesize the outcome (and it is therefore exploratory in nature) or because the duration of the evaluation does not allow sufficient time to assess the hypothesized change.⁵

DESIGN AND ANALYTIC PROCEDURES

A brief overview of the study design methods and analytic procedures for the Impact Evaluation is provided below; detailed procedures can be found in Appendix A.

EVALUATION DESIGN

The Evaluation Team's approach was to assess the impact of BEAM by means of a RCT, where eligible individuals were randomly assigned to one of two conditions: (1) the offer to participate in the innovative set of BEAM services (GCP), or (2) the offer to participate in AJC services as usual (CCS). The Impact Study contrasted the employment and educational outcomes of individuals randomly assigned to the treatment condition (GCP) with those of individuals assigned to receive services that they would typically be offered by the AJC (control condition). The purpose of the study was to determine BEAM's

⁵ Though the Evaluation Team makes the distinction between primary and exploratory outcomes, the estimates produced in response to the primary and exploratory research questions are based on the same rigorous methods and therefore are both causal in interpretation.

efficacy in improving these outcomes for postsecondary dropouts. The unit of assignment was the individual participant; the unit of analysis was the individual participant.

RECRUITMENT AND ENROLLMENT

The target population for BEAM was individuals 19 years or older who had previously attempted a postsecondary education training program, discontinued enrollment prior to completing the program, and were residents of the nine-county region where the BEAM intervention was offered.⁶ To be eligible to participate in the Impact Study, a customer had to be 19-years-old, or older, a postsecondary drop out, and a resident of one of the nine counties within the implementation region. Customers also had to give consent to participate in the study, be willing to return to an AJC to receive services, register with the AJC, not be a veteran, and not be currently enrolled in postsecondary training.

Outreach Coordinators (OCs) met in person with individuals to determine whether they were eligible to participate in the Impact Study, per the criteria outlined above. Eligible participants who provided consent were randomly assigned to the treatment or control condition at a 4:1 ratio (treatment to control), such that 80% of enrolled participants were offered the treatment condition (GCP) and 20% of participants were offered the control condition (CCS).

Table 1 presents the total number of individuals who were randomized in the Impact Study in each of the four WDB regions, overall and by treatment condition. A total of 365 participants were enrolled in the Impact Study, 289 in the treatment group and 76 in the control group.

WDB Region	Number Randomized	Treatment	Control
НМО	111	89	22
CDO	90	70	20
B-T	83	65	18
Tompkins	81	65	16
Total	365	289	76

 Table 1. Number of Participants Randomized, by WDB Region and Condition

DATA

Individual-level outcome data, background characteristics, contextual economic data, and fidelity data were collected from four main sources for the Impact Study: New York State Department of Labor (NYSDOL), the National Student Clearinghouse (NSC), BEAM OCs, and the U.S. Bureau of Labor Statistics (BLS). The Evaluation Team established a data sharing agreement with NYSDOL, which is the source of data for participant background characteristics and employment outcomes. A data sharing agreement was also established with the NSC, which is the source of data for educational outcomes. OCs regularly submitted data collected and entered during screening and enrollment procedures on eligibility criteria and the results of randomization for each individual enrolled. The Evaluation Team also collected data that are publicly available through the BLS on regional economic indicators such as local unemployment rates and hourly earnings.

TREATMENT CONTRAST

As outlined in the Program Design section of this report, BEAM was a program model that employed a core network of staff (Project Director, Communications Coordinator, and OCs) to deliver individualized

⁶ BEAM was offered in Broome, Chenango, Delaware, Herkimer, Madison, Oneida, Otsego, Tioga, and Tompkins Counties in New York.

assistance to reenroll disconnected youth into education and employment opportunities that went above and beyond what was traditionally provided at AJCs prior to its inception. Individuals enrolled in GCP (treatment) were offered individualized support from OCs but could have also benefited from traditional services and funding under other federal programs available to AJC customers. The control group (CCS) was only offered the traditional suite of services. The Impact Study aimed to assess the educational and employment impact of being offered GCP (treatment) relative to being offered CCS (control). The treatment contrast is therefore all the services offered by GCP that are not offered by CCS. The control condition can be understood as a business-as-usual (BAU) contrast. Using a BAU contrast rather than something else (e.g., a no-intervention control) is useful because it contrasts the experimental intervention (in this case GCP) with current practice and provides estimates that represent the improvement (or decline) over current practice. The impact estimates produced by this study, therefore, are best conceived as the average "improvement" in outcomes that are attributable to GCP above and beyond what is being achieved by current CCS practice.^{7, 8}

GUIDED CAREER PIPELINE

Individuals who were enrolled in the treatment condition were offered the opportunity to receive individual educational and employment assistance from OCs, as well as traditional services and funding typically available to any AJC customer. Opportunities, facilitated by the OCs, that would have been available only to treatment participants include (1) one-on-one navigation, advocacy, and customized assistance through the entire process from reenrollment to employment; (2) Individual Academic Plan development, revisions, goal planning, and career counseling; (3) postsecondary education financial aid support (e.g., counseling, navigation of institutions' financial aid offices, one-on-one assistance with completing applications, navigation of default loan rehabilitation); (4) regular check-ins with OC (or partner staff), ranging from daily to monthly; and (5) referrals to additional community resources.

In addition to the opportunities outlined above, GCP participants could have received the following services, which are typically available to AJC customers: job readiness assessments, job counseling, skill assessment, career planning, referrals to and funding for training, computer and Internet access for resume development and job searching, printing services, and library access.

CAREER CENTER SERVICES

Individuals who were enrolled in the CCS condition of the Impact Study were referred to an AJC staff person who then offered the individual educational enrollment and employment assistance that is identical to what they would have been offered at each of the AJCs prior to the introduction of BEAM. For study enrollments that were conducted on-site at the AJC, the OC would offer an in-person referral to AJC staff, and for individuals who were enrolled off-site (e.g., at a community partner site), the OC would provide the individual with the AJC staff's contact information and inform the staff that a referral had been made. Specific services that were offered to CCS participants include job readiness

⁷ In other words, a null outcome does not represent a lack of improvement in participant outcomes, but rather no relative improvement over what is currently being achieved by BAU at the AJCs. A positive impact represents a desirable improvement over current practice and a negative impact represents an undesirable decline over current practice.

⁸ Note that the Implementation Evaluation section of the full evaluation report discusses the innovative intervention as BEAM, which is the program model that utilizes OCs to provide one-on-one navigation support that is not available through the standard AJC service model. The Evaluation Team does not discuss BEAM as the treatment intervention in the Impact Evaluation section because participants enrolled in the treatment condition could have also benefited from BAU services on top of those offered by BEAM staff. GCP is the package of services that OCs delivered that went above and beyond what would have been available to AJC customers. The GCP package was delivered by BEAM staff and is a core component of BEAM, but is not the whole extent of the BEAM experience. The Impact Study's treatment contrast attempts to isolate the impact of the additional services available to GCP customers that were not available to CCS customers.

assessments, job counseling, skill assessment, career planning, referrals to and funding for training, computer and Internet access for resume development and job searching, printing services, and library access. The BAU condition naturally varied by AJC, county, WDB region, and each participant's educational and career goals – there were 11 AJCs, across 9 counties, within 4 WDBs, and participants entered the study at varying degrees of financial stability – but participants assigned to the control condition received the same core structure of services under AJC funding streams, such as the Workforce Innovation and Opportunity Act (WIOA) and the Wagner-Peyser Act.⁹

INTENT-TO-TREAT FRAMEWORK

The principal aim of this study is to determine whether offering the GCP program to participants improves their educational and economic outcomes. This is done within the intent-to-treat (ITT) framework, which means that the analysis aims to include all the participants enrolled in the Impact Study, regardless of the dosage or exposure to program components. Although this approach can seem obtuse because it fails to account for the variation in participants' actual exposure, researchers adopt it because it provides the most unbiased estimate of program impact. An ITT estimate minimizes the potentially biased post-enrollment self-selection that motivates some people to engage more and others to engage less with the intervention. This estimate also has the added advantage of providing a more realistic estimate of the predicted impact of the program because it factors in the variation of exposure into the estimate, rather than controlling for it statistically.

ANALYSIS

The impact of the GCP program is estimated with a regression equation that models outcomes as a function of treatment status, the baseline measure of the outcome variable (or a proxy), and other covariates. Although a straight difference-of-means/proportion approach would provide unbiased estimates of the effect of the treatment intervention, a statistical model that includes covariates is preferred because it increases the precision of the impact estimates. Maximizing precision is particularly necessary in this study because of the small sample size. An ordinary least squares (OLS) model was used to estimate all outcomes (using Stata 15).

BENCHMARK SAMPLE

As discussed in Appendix A, the Evaluation Team initially proposed to enroll 1,800 participants at a 4:1 ratio (treatment to control) by the end of December 2017 and assess outcomes at the end of December 2018.¹⁰ Challenges in the enrollment process resulted in an analytic sample that was substantially smaller. From December 2015 through March 2018, 365 participants were enrolled in the study at a 4:1 ratio.¹¹

The Evaluation Team received outcome data from the NYSDOL and NSC for all participants enrolled in the study. Therefore, the analytic sample and the ITT sample are identical.¹² Nominally, this suggests that the analytic sample is complete and should not be biased by post randomization selection. In

⁹ For the sake of simplicity, any mention of WIOA programming or funding streams with regard to the Impact Study implies the Workforce Investment Act (WIA) equivalent when considering data prior to the implementation of WIOA in July 2016.

 $^{^{\}rm 10}$ This would have allowed a full year of follow-up for all participants in the sample.

¹¹ This sample is substantially smaller than was proposed based on a power analysis and is considered by the evaluators to be under-powered. An a priori power analysis with standard assumptions, conservative expectations, and a total analytic sample size of 1,800 participants (1,440 treatment and 360 comparison) indicated that the minimal detectable effect size (MDES) for an outcome of this sample size would be d = 0.17. ¹² See Appendix A for a discussion of the ITT framework.

actuality, the completeness is an artifact of the structure of the data. This may hide attrition and bias impact estimates.¹³

As a result of the enrollment challenges, The Evaluation Team decided to extend the benchmark study for an additional quarter (January through March 2018) for the purpose of increasing sample size and statistical power. As a consequence, outcomes had to be assessed at the end of three quarters postenrollment, rather than the full year, as specified in the design plan. Impact estimates from the benchmark sample are referred to as the "full sample" throughout the Findings sections.

FULL-YEAR SUBSAMPLE

Primary results presented below are statistically indeterminate for all outcomes. Given the consistency of findings, however, the Evaluation Team reasoned that much of this was motivated by limited sample size. In response to this, and to test the secondary hypothesis (of possible impacts hiding behind small-sample null results), the secondary study examines a subsample of individuals who had at least a full year after enrollment in BEAM to achieve outcomes. Therefore, throughout the Findings section, benchmark analytic results are presented alongside additional secondary study results that restrict the analytic sample to those participants who were enrolled in the Impact Study by the end of December 2017. This additional exploratory evidence is offered alongside the benchmark findings to improve the quality of the analysis and provide BEAM leadership with constructive information about the efficacy of outcomes in the context of the confirmatory findings. Analysis and interpretive discussion are offered where secondary (full-year subsample) study results provide additional insight. Otherwise, the findings are simply reported, and readers are referred to Appendix A for further details.

Although the total number of participants is smaller than those in the benchmark confirmatory study (*n* = 341), it was the Evaluation Team's hypothesis that this group should evince stronger or more pronounced estimates because they were a group with more time to realize the expected effects. This is essentially a test of whether greater exposure to the intervention leads to comparatively improved outcomes. If the program is working as intended, this should indeed be the case. And, even if the results are not significant (they are not), some evidence of more positive (in the hypothesized direction) effects is indicative of promise – even if it is not confirmatory evidence of program impact. The results for this smaller sample are presented not as an alternative to the confirmatory findings but as an addendum. They are considered secondary because they refer to a subgroup of the full study sample, and because this analysis was developed after the decision to extend enrollment through March 2018 was made. Results from these secondary analyses are referred to as "subsample" findings throughout the Findings sections.

UNEMPLOYED SUBSAMPLE

In addition to the full-year subsample analysis discussed above, the Evaluation Team conducted an additional exploratory analysis on a subsample of participants who were unemployed during the quarter that directly preceded the quarter of study entry to determine if there was any variation in observed impact estimates compared with findings from the benchmark sample. Results from this analysis are

¹³ As explained in Appendix A, both the NSC and NYSDOL do not differentiate between data that are missing and data that are "true" zeros; in other words the Evaluation Team cannot determine whether someone did not achieve the outcome because their data were unavailable or because they did not enter training or become employed. In the event that data were missing at random for both the treatment and control participants, the threat of bias would be minimal. If, however, there was an imbalance in missing data between the two groups, this would have biased results. For example, were more data missing for the treatment than control group, this would attenuate program impact estimates, whereas if more data were missing for the control group, this would artificially inflate impact estimates.

briefly discussed in the discussion of employment outcomes. These results are considered exploratory because they refer to a small subgroup of the full study sample and because this analysis was developed after the data became available. The details of these analyses are provided in Table A9 of Appendix A but are not discussed in the main Findings section of this report.

FINDINGS

OVERVIEW

Results from the Impact Study are presented below. A discussion of the background characteristics of the benchmark sample and the full-year subsample is presented first. Results are presented individually for each of the six research questions.

Overall, benchmark findings indicate that the GCP program did not have a statistically discernible impact on participants' educational and employment outcomes when compared with participants in the CCS program. Within the educational domain, benchmark estimates (of the mean impact) suggest that the treatment group appears to be enrolling, persisting, and completing postsecondary programs at higher rates than the control group – as was hypothesized – although the magnitude of the differences is too small and too varied for the Evaluation Team to be confident that the mean effect is statistically distinguishable from no effect. Additional analysis of the full-year subsample provides some evidence to suggest that the effect of the GCP program on participants' educational outcomes does increase with time though estimates are not significant.

Within the employment domain, benchmark findings indicate that the GCP program had no detectable impact on participant outcomes (employment, wages, employment persistence). Secondary analyses conducted on a subgroup of participants who had more time to achieve outcomes corroborate these findings. Exploratory findings on a smaller subgroup of participants who were unemployed at the time of study entry suggest, however, that the GCP program could potentially benefit those individuals who were not employed upon program enrollment above and beyond the CCS program. A detailed overview of the Impact Study findings and an interpretive discussion of these results are provided below.

SAMPLE CHARACTERISTICS

Table 2 presents the background characteristics of the combined (both treatment and control) Impact Study benchmark sample, as well as the full-year subsample.¹⁴ As discussed in Appendix A, the benchmark analytic sample is identical to the ITT sample, due to the nature of the data used to assess educational and employment outcomes. NSC and NYSDOL Unemployment Insurance (UI) wage data do not provide a clear avenue to distinguish between data that are missing and those that are "true" zeros. As such, both overall and differential attrition are nominally set at zero, but these statistics are optimistic.¹⁵

Participants in the benchmark sample are on average, between 30 and 31 years old and majority female (64%). Most identify as White (67%) or Black (26%). The remaining identify as Multiracial (6%) or some other race (2%), and 10% identify as Hispanic or Latino. A little over one-half (54%) had low income status according to NYSDOL guidance at the time that they enrolled into the study; 6% were receiving Supplemental Security Income (SSI) or Social Security Disability Insurance (SSDI) benefits, and 11% were

¹⁴ Baseline balance statistics between the benchmark and full-year subsample treatment and control groups are provided in the Baseline Equivalence section of Appendix A.

¹⁵ See footnote 43

receiving Temporary Assistance for Needy Families (TANF) benefits at the time of enrollment in the study. Just under one third (28%) of study participants had some kind of disability (as defined by NYSDOL standards) at the time of study enrollment. The majority (76%) were registered under the WIOA adult funding stream at the time they enrolled in the study and, to a lesser extent, dislocated worker funding (27%), whereas a small minority were youth funding recipients (7%). A little over one third (36%) of study participants identified as having a high school or equivalent education level at the time of study enrollment, whereas half (50%) had some college, but no degree, and 13% had an Associate's Degree or higher.¹⁶ About 59% of study participants were employed and earned an average of \$2,524 during the quarter that immediately preceded study enrollment.

Compared with the benchmark sample of participants, the full-year subsample of participants enrolled in the study by December 2017 are overall very similar. In addition to the overall composition of the benchmark sample and the subsample, baseline balance statistics between the treatment and control groups of both samples are provided in Appendix A.¹⁷

¹⁶ The astute reader will recognize that the data reflecting highest level of education at study enrollment potentially contradicts the Impact Study eligibility criteria outlined in the Design Summary section, namely that an individual must have attempted postsecondary training, but dropped out prior to obtaining a degree. There are a few possible explanations for this discrepancy. First, the eligibility criteria do not limit an individual who may have obtained an Associate's or Bachelor's Degree, attempted a higher degree, but failed to complete it from joining the Impact Study. Second, the workforce system provides postsecondary training opportunities through its network of Board of Cooperative Educational Services (BOCES), but which are not accredited colleges or universities. The Evaluation Team agreed that this level of postsecondary training would satisfy the criteria that an individual enroll in postsecondary training but dropped out before completing a degree or program. ¹⁷ As discussed in Appendix A, a well-executed randomization procedure does not guarantee balance on any given characteristic; what it does guarantee is that these features will be independent of the assignment to treatment or comparison condition. As such, the Evaluation Team presents baseline equivalence statistics as a descriptive feature and a means to identify any irregularities in Appendix A.

BEAM IMPACT EVALUATION

Table 2. Background Characteristics of Impact Study Participants¹⁸

	Benchma	rk Sample	Full-year S	Subsample		
	Number		Number			
Characteristic	Reported	Statistic	Reported	Statistic		
Age	(<i>n</i> = 365)		(<i>n</i> = 341)			
Mean age in years at enrollment	365	30.6	341	30.0		
Gender	(<i>n</i> = 343)		(<i>n</i> = 319)			
Female	220	64.1%	205	64.3%		
Race	(<i>n</i> = 326)		(<i>n</i> = 304)			
White	218	66.9%	206	67.8%		
Black	84	25.8%	75	24.7%		
Multiracial	19	5.8%	18	5.9%		
Other	5	1.5%	5	1.6%		
Ethnicity	(<i>n</i> = 284)		(<i>n</i> = 262)			
Hispanic/Latino	29	10.2%	27	10.3%		
Low income status	(<i>n</i> = 365)		(<i>n</i> = 341)			
Yes	197	54.0%	183	53.7%		
Receives SSI or SSDI	(<i>n</i> = 365)		(<i>n</i> = 341)			
Yes	23	6.3%	20	5.9%		
Receives TANF	(<i>n</i> = 365)		(<i>n</i> = 341)			
Yes	40	11.0%	38	11.1%		
Individual with a disability	(<i>n</i> = 327)		(<i>n</i> = 304)			
Yes	92	28.1%	86	28.3%		
WIOA recipient ¹⁹	(n = 365)	76 40/	(n = 341)	76.00/		
Adult	2/9	76.4%	262	/6.8%		
Youth	2/	7.4%	26	7.6%		
Dislocated worker	9/ (n = 265)	20.0%	90 (n = 241)	26.4%		
	(77 = 305)	26.20/	(1 = 341)	25.00/		
As Diploma or equivalent	132	50.2%	172	50.7%		
Associate's Degree or higher	184	13.4%	175	13 5%		
Proportion employed pre-enrollment	(n - 365)	13.470	(n - 3/1)	15.570		
1 st quarter	216	59.2%	197	57.8%		
2 nd quarter	210	60.8%	207	60.7%		
3 rd quarter	213	58.4%	200	58.7%		
4 th guarter	214	58.6%	201	58.9%		
Average proportion employed	365	59.2%	341	59.0%		
Mean quarterly wages pre-enrollment	(<i>n</i> = 365)		(<i>n</i> = 341)			
1 st quarter	365	\$2,524.27	341	\$2,502.75		
2 nd quarter	365	\$2,733.85	341	\$2,683.53		
3 rd quarter	365	\$2,778.75	341	\$2,767.43		
4 th quarter	365	\$2,898.87	341	\$2,893.81		
Average wages	365	\$2,733.93	341	\$2,711.88		
Regional labor statistics	(<i>n</i> = 365)		(<i>n</i> = 341)			
Mean pre-enrollment unemployment rate	365	5.4%	341	5.4%		
Mean pre-enrollment labor force	365	20,771	341	20,847		
Mean pre-enrollment average hourly earnings	365	\$29.60	341	\$29.55		
Mean pre-enrollment average hours worked in a week	365	33.5	341	33.5		
Mean pre-enrollment minimum wage	365	\$9.47	9.47 341			

¹⁸ Sample size is reported for each characteristic reported in Table 2. In some cases, data were not available for a characteristic for the entire study sample (e.g., gender, race, ethnicity, disability status). Missing baseline characteristic data were imputed when included in benchmark analytic models. No outcome data were imputed.

¹⁹ Individuals could be registered under more than one WIOA funding stream and as such the proportions reported under this section exceed 100%.

KEY FINDINGS

In the section below, results from the benchmark and secondary analysis are presented for each research question. The benchmark analysis includes the full ITT sample of participants who were enrolled in the study by the end of March 2018.²⁰ Detailed benchmark impact estimates are presented in Table A7 of Appendix A. Full-year subgroup analyses (for the subgroup of participants who were enrolled by the end of 2017) are presented alongside the benchmark findings in the figures below. Detailed full-year subsample impact estimates are presented in Table A8 of Appendix A. An interpretive discussion of these results is provided in the Discussion section.

RESEARCH QUESTION 1: IMPACT ON POSTSECONDARY ENROLLMENT

Benchmark statistical estimates for Research Question 1 indicate that the GCP intervention has no statistically detectable effect on participants' likelihood of enrolling in a postsecondary education program after enrolling in BEAM. Model estimates presented in Table A7 and graphically illustrated in Figure 1 below show that after nine months GCP participants were, on average, 3% more likely to enroll in postsecondary training than participants offered CCS. While these estimates show that members of the treatment group appear to be enrolling in postsecondary education at a slightly higher rate than those in the control group – as was hypothesized – the estimated mean difference is too slight and there is too much variation in the estimate for the Evaluation Team to be confident that the difference is greater than zero.





To assess whether the effects of the GCP program were more pronounced for individuals who were enrolled in the study for a longer period of time, the Evaluation Team looked at the subsample of participants who enrolled by December 2017. As shown in the second set of bars in Figure 2, this

²⁰ The BEAM Project Director and the Evaluation Team decided to extend enrollment past the timeline originally outlined in the EDR, which would have ended study enrollment after December 31, 2017.

²¹ Figure 3 presents the regression adjusted means for the percentage of participants who enrolled in postsecondary training within nine months (full sample) and one year (full-year subsample) after study entry.

appears to be the case. In this subsample, GCP participants are 8% more likely to enroll in postsecondary training than their CCS counterparts.²²

Although results are not statistically significant, with an additional three months of time, the magnitude of the estimated average difference between GCP and CCS postsecondary training enrollment grows, and the effect size triples from 0.1 (benchmark) to 0.3 (subsample).²³ For those participants who did enroll in a postsecondary education program, the average time to enrollment after study entry was between seven and eight months.²⁴

RESEARCH QUESTION 2: IMPACT ON POSTSECONDARY PERSISTENCE

Benchmark findings indicate that GCP does not have an impact on participants' persistence (defined as the number of consecutive semesters enrolled) in postsecondary education program. Figure 2 illustrates this by graphically depicting the regression-adjusted mean number of consecutively enrolled semesters for each treatment group. The figure shows that, on average, GCP participants were enrolled for just one tenth of a semester longer than CCS participants.





A descriptive representation of the number of consecutive semesters that participants enrolled in postsecondary education is presented in Figure 3 below. The numbers on the *x*-axis represent the number of consecutive semesters enrolled and the height of each bar illustrates the proportion of participants who achieved that enrollment level. The figure graphically shows that most participants enrolled in zero consecutive semesters and that the remaining minority attended slightly more semesters in diminishing quantity. Comparatively speaking, the distribution looks the same for both the GCP and CCS groups.

²² Detailed impact results from the full-year subsample analysis can be found in Table A8 in Appendix A.

²³ Effect size is defined as a standardized measure of the magnitude of program effect. It represents the estimated impact of the program on that outcome in standard deviation units. Although it is conventional to consider any effect less than .20 as small, that is an overly generalized interpretation of Cohen's effect size guidelines. Any determination of size must consider the nature of the intervention (e.g. cost, intensity, etc.), the noise inherent in the outcome measure, and the outcome itself.

²⁴ A series of sensitivity tests were conducted to test the extent to which the benchmark findings were robust to different analytic decisions and assumptions. The impact estimates from these additional tests (presented in Table B2 of Appendix B) are consistent with benchmark results.
²⁵ Figure 2 presents the regression-adjusted mean number of consecutive semesters participants were enrolled in postsecondary training.

BEAM IMPACT EVALUATION

Figure 3. Percentage of Benchmark Sample Participants Who Were Enrolled in Postsecondary Training by Number of Consecutive Semesters (n = 365)²⁶



The outcome for Research Question 2 is calculated as the number of consecutive semesters enrolled in postsecondary training through the end of the study (December 2018). Participants in the benchmark sample would have had a follow-up period that ranged from a minimum of 9 months (i.e., those who entered at the end of study enrollment) to a maximum of 36 months (i.e., those who entered at the beginning of study enrollment). The secondary study narrows this range from 12 to 36 months, which does not extend the duration of time but does increase the average time allowed. As shown in Figure 3 above, this subsample adjustment has little detectable effect. Subsample findings are consistent with benchmark findings – GCP participants were enrolled for just a tenth of a semester longer than CCS participants. Results remain not significant and the magnitude of the estimated effects are substantively identical in terms of effect size. Detailed results can be found in Appendix A.²⁷

RESEARCH QUESTION 3: IMPACT ON DEGREE ATTAINMENT

The benchmark analysis indicates that the GCP program did not have a detectable impact on participants' likelihood of graduating from a postsecondary education program by the end of the study period. Model estimates presented in Table A7 and graphically illustrated in Figure 4 below show that by the end of the study (December 2018), GCP participants were, on average, less than 1% more likely to graduate from postsecondary training than participants offered CCS. Again, the estimated mean difference is too small and uncertain to permit a statistical inference of difference.

²⁶ Figure 3 presents the unadjusted percentage of participants who were enrolled in postsecondary training for consecutive semesters.
²⁷ A series of sensitivity tests were conducted to test the extent to which the benchmark findings were robust to different analytic decisions and assumptions. The impact estimates from these additional tests (presented in Table B3 of Appendix B) are consistent with benchmark results.

BEAM IMPACT EVALUATION



Figure 4. Percentage of Participants Who Graduated From Postsecondary Training²⁸

As with Research Question 2, individual follow-up windows would have ranged from 9 to 36 months for participants in the benchmark sample. In the full-year subsample analysis, the analytic sample was restricted to those who had 12 to 36 months to realize training outcomes. Findings from the secondary subsample are consistent with benchmark findings – GCP participants were about 1% more likely to graduate from their postsecondary program than CCS participants by the end of the study. The standardized effect size for both analyses is 0.2, which means that the estimated magnitude of the effect is equally small (relative to the observed variation in the outcome).^{29, 30}

RESEARCH QUESTION 4: IMPACT ON EMPLOYMENT

Benchmark findings indicate that the GCP program had no effect on participants' likelihood of becoming employed three quarters after study entry. Statistical estimates (presented in Table A7 of Appendix A) demonstrate that the GCP program did not have a detectable effect on employment outcomes. At the end of the 3rd quarter, following enrollment in the study, 69% of treatment participants and 70% of control participants were employed for at least one quarter post-enrollment. As shown in Table A7, the estimated difference between the two groups' likelihood of becoming employed is effectively zero.

Figure 5 presents the unadjusted proportion of individuals who were employed during four preenrollment quarters, quarter of study entry, and three post-enrollment quarters, by treatment condition.

²⁸ Figure 4 presents the regression-adjusted means for the percentage of participants who obtained a degree or certificate from postsecondary training by the end of the study.

²⁹ In total, only two participants from the benchmark control group and nine from the treatment group were identified as having graduated from their postsecondary education program. Of those who did graduate, they did so, on average, between 17 and 18 months after study entry. Two participants earned a professional certificate, eight earned an Associate's Degree, and one completed a Bachelor's Degree.
³⁰ In addition to the secondary analysis presented here, the Evaluation Team conducted a series of sensitivity analyses to test the robustness of the benchmark approach, all of which are consistent with the benchmark findings and presented in Table B4 of Appendix B.



Figure 5. Percentage of Participants Employed During Pre- and Post-Enrollment Quarters (n = 365)³¹

Both lines demonstrate that post-program employment for both conditions is perhaps slightly better than it was at the time of enrollment, but not dramatically so. They also demonstrate how similar the two groups' employment trends are before and after study enrollment. The uncertainty of difference is illustrated by the crisscrossing of the two lines post enrollment; this vacillation hints at the absence of a detectable effect between the two conditions and demonstrates the uncertainty of any consequent impact estimates. At each time point there are small differences between the two groups (e.g., 55% of GCP participants were employed during the first full quarter after study entry, compared with 51% in the CCS group); however those differences are in flux (e.g., a positive difference at post quarter 1 becomes a negative difference at post quarter 2). This inconstancy could be motivated by actual differences, but with the small comparison sample, it is likely imprecision driven by measurement error and the natural variation in human experience.

Results for the full-year subsample analysis are substantively identical to the benchmark findings. For details, see Table A8 and Figure A3 in Appendix A.³²

RESEARCH QUESTION 5: IMPACT ON QUARTERLY WAGES

Benchmark findings suggest that the GCP program had no detectable effect on participants' quarterly earnings three quarters after study entry. Statistical estimates are presented in Table A7 of Appendix A and indicate that the GCP program did not have a significant impact on wages earned during the 3rd quarter after study entry. On average, treatment group participants earned a total of \$2,862 during the 3rd quarter, whereas control group participants earned a total of \$2,736, a difference of \$126. While the treatment group in the benchmark sample earned more money, the mean difference was too small and the variability too large to be confident that the mean estimate is reliably greater than zero.

³¹ Figure 5 presents the unadjusted means for the percentage of participants who were employed during each of the four quarters prior to study enrollment, the quarter of study enrollment, and up to three quarters after study enrollment. Analytic sample size is 365 for all time points.

³² A series of sensitivity tests were conducted to test the extent to which the benchmark findings were robust to different analytic decisions and assumptions. The impact estimates from these additional tests (presented in Table B5 of Appendix B) are consistent with benchmark results.

The next figure illustrates that indeterminacy well. Figure 6 presents comparative unadjusted average quarterly wages for both groups across four pre-enrollment quarters, quarter of study entry, and three post-enrollment quarters. Again, the figure demonstrates small and wavering differences in participants' estimated wages following enrollment in the study. This time, however, it appears that that there is noticeable improvement for both groups after study enrollment. Wages for both groups improve in an almost linear fashion, immediately after enrolling in the GCP and CCS programs. This suggests that something common to both programs, or common to the experience of both groups is improving wages for both.





Again, full-year subsample results are substantively identical to the benchmark findings. For details, see Table A8 and Figure A4 in Appendix A.³⁴

RESEARCH QUESTION 6: IMPACT ON EMPLOYMENT PERSISTENCE

Benchmark findings suggest that the GCP intervention had no significant effect on participants' likelihood of staying employed after enrolling in BEAM. Model estimates presented in Table A7 and illustrated in Figure 7 below indicate that three quarters following enrollment there was no statistically significant difference between the two groups' persistence outcomes; both GCP and CCS participants were employed for 1.7 quarters (out of 3).

³³ Figure 6 presents the unadjusted means for the quarterly earned wages during each of the four quarters prior to study enrollment, the quarter of study enrollment, and three quarters after study enrollment. Analytic sample size is 365 for all time points.

³⁴ The results of a series of sensitivity analyses on the benchmark approach, all of which are consistent with benchmark findings, are presented in Table B6 of Appendix B.

BEAM IMPACT EVALUATION



Figure 7. Average Number of Quarters Employed, by Treatment Condition³⁵

Findings for the full-year subsample are consistent with those from the benchmark sample. As illustrated in the lower panel of Figure 7, both groups have improved employment outcomes, but the difference between the two groups is effectively unchanged. Differences are not statistically significant and are small in magnitude (effect size of zero for both). For details, see Appendix A.³⁶

DISCUSSION

POSTSECONDARY ENROLLMENT, PERSISTENCE, AND COMPLETION

The benchmark analysis finds that the GCP intervention under BEAM did not have a detectable impact on primary or exploratory outcomes. Education outcomes for GCP customers were statistically indistinguishable from CCS customers. However, and with particular emphasis on Research Question 1, the Evaluation Team has confidence that the statistical findings observed here are primarily motivated by sample size. This assessment is based on three features of the study. First, the final analytical sample size was small – much smaller than specified in the EDR and smaller than an a priori power analysis indicated would be sufficient to detect an effect.³⁷ Second, benchmark findings, which were modest in magnitude, were consistently in the hypothesized direction under the educational outcome domain. Third, secondary (full-year subsample) analyses, which the Evaluation Team undertook because of the decision to extend the enrollment period, corroborate the benchmark findings in the sense that they are (not significant but) all in the hypothesized direction, and intimate that effects tend to increase with increased dosage or time.³⁸

³⁵ Figure 7 presents the regression-adjusted means for the number of quarters employed after study entry out of three quarters (full sample) and out of four quarters (subsample).

³⁶ The results of a series of sensitivity analyses on the benchmark approach, all of which are consistent with benchmark findings, are presented in Table B7 of Appendix B.

 $^{^{37}}$ The EDR originally proposed to enroll 1,800 participants into the study at a 4:1 ratio where 1,440 would be enrolled in the treatment group and 360 into the control group. An a priori power analysis with standard assumptions, conservative expectations, and a total analytic sample size of 1,800 participants (1,440 treatment and 360 comparison) indicated that the minimal detectable effect size (MDES) for an outcome of this sample size would be d = 0.17.

³⁸While these additional analyses may be promising, they should be understood as being exploratory in nature. The analyses are based on a purposive subsampling of the full ITT sample.

In terms of postsecondary education enrollment (Research Question 1), the benchmark findings provide no statistical evidence to suggest the GCP intervention has an impact on participants' likelihood of enrolling in a program within nine months. Estimates are in the direction that BEAM leadership expected, but impacts are modest in magnitude and not statistically significant. Secondary (full-year subsample) analyses suggest that the program may be having a greater effect than is captured by the confirmatory study. Findings from the Implementation Study further suggest that educational results may have been even greater, had the evaluation been able to measure outcomes over a longer period. Clients faced several barriers to postsecondary enrollment after study entry, including resolving defaulted student loans, choosing an appropriate program, navigating the education system, and lack of funds to pay for education costs. In other words, enrollment in college is not just a simple matter of completing application materials. The process can take months or longer, depending on the time of year and the circumstances of the client. Further, these barriers likely do not just inhibit enrollment, but they could also attenuate enrollment outcomes in the short term.

The Evaluation Team conducted an additional analysis to test this. Specifically, the additional analyses allowed participants to have longer (a full year) to enroll in a postsecondary program. Findings confirm the expectation that GCP enrollment results improve with time. In particular, findings indicated that the comparative rate at which GCP participants enroll in a program versus CCS participants increases from 3% at the nine-month time point to 8% at the one-year mark. Considering that the average time to postsecondary enrollment was between seven and eight months, and that the standardized effect size tripled from 0.1 for the nine-month period to 0.3 for the one-year period, the Evaluation Team believes that, all else being equal, had the BEAM Team been able to enroll a larger number of participants into the Impact Study within the original time frame (end of 2017), results may have been statistically significant.

Postsecondary persistence (consecutive semesters enrolled) and completion outcomes are more modest, but given that these are more distal outcomes (than enrollment) the evaluation may simply have not allowed enough time for participants to realize these outcomes fully.³⁹ On average, participants were engaged in the Impact Study for a period of 22 months. Considering that it took customers an average of seven to eight months to enroll in postsecondary training, it should not, perhaps, be surprising that the relative difference in persistence and completion outcomes were muted.

An additional qualifier here is that the data that were used to measure postsecondary outcomes for the Impact Study are not comprehensive. Even in the catchment area of the BEAM region, the data reported to the NSC do not include all postsecondary educational institutions or certification programs that study participants could select. By way of example, the Implementation Study identifies several cases where participants pursued industry-recognized certifications at institutions that were not reflected in the data made available to the Evaluation Team. The NSC data included only accredited two- or four-year private and public colleges and universities. Programs such as those offered through local Board of Cooperative Educational Services (BOCES) or industry certifications offered on the job, while relevant to BEAM's overarching program model and goals, would not have been included in the program completion data available to the Evaluation Team.

³⁹ This is an external constraint imposed on the evaluation by the scope of the grant.

EMPLOYMENT AND WAGES

The Impact Study assessed GCP program impact on employment outcomes, including employment, average quarterly wages, and employment persistence, compared with the business-as-usual (BAU) condition. Benchmark results indicate that the GCP intervention did not have any detectable impact on employment outcomes through the third full quarter after study entry. For the most part, the impact estimates are either close to zero, or suggest a potential negative effect, though none are statistically significant. Secondary analyses on the subsample of participants who had a full year (4 quarters) of follow-up are similarly null.

In addition to the benchmark and full-year subsample analyses described above, the Evaluation Team conducted a third, exploratory subgroup analysis that limited the analytic sample to individuals who were unemployed during the quarter that directly preceded study entry. Results from the exploratory analyses conducted on this unemployed subgroup were not significant as well; however, there is evidence of promise here. As shown in Table A7 of Appendix A, benchmark impact estimates and their standardized effect sizes were close to zero for all three outcomes under the employment domain. When the sample was restricted to the subgroup of participants who were unemployed directly prior to study entry (n = 149), impact estimates increased, were consistently positive for all three outcomes (though remained not significant), and standardized effect sizes increased to 0.3 for Research Questions 4 (employed) and 6 (employment persistence).⁴⁰ Although the results of these unemployed subgroup analyses are exploratory (e.g., not confirmatory) in nature, they do provide some additional context for GCP's impact on participants' employment outcomes, namely that the program may benefit individuals who are unemployed at entry more so compared with the general study population.

The primary goal of BEAM was to connect postsecondary dropouts with education and training opportunities that would allow them to obtain industry-recognized credentials, and eventually higher paying jobs. The time frame of the impact study was limited by the scope of the funding grant; outcomes could be assessed for up to four quarters but the long-term impact of the GCP intervention was always beyond the scope of this analysis. Three to four quarters may simply have not been enough time for participants to realize improved labor market outcomes. The Evaluation Team expressed this hypothetical uncertainty in the EDR; it was not clear whether participants who were selected to receive GCP, which included one-on-one navigation through the education system, and were expected to have higher rates of postsecondary enrollment, would have time to realize employment outcomes, especially considering that participants assigned to the CCS condition would be expected to more immediately pursue employment opportunities. It was for these reasons that labor market outcome questions were not identified as confirmatory.

CONCLUSIONS AND STUDY LIMITATIONS

The benchmark Impact Study findings indicate that the GCP program did not have any discernible impact on participant educational or employment outcomes, namely postsecondary enrollment, persistence, and completion or likelihood of employment, average wages, and quarters employed. The results indicate that the participants who received the GCP intervention were equally as likely to enroll in a postsecondary education program and become employed as the participants who were offered the AJC BAU experience. There is some exploratory (e.g., not confirmatory) evidence to suggest that given a

⁴⁰ See Table A9 in Appendix A for detailed impact estimates.

longer study window, GCP participants would have been more likely to enroll in a postsecondary program compared with CCS participants.

A well-executed RCT offers the most internally valid estimates of a program's impact. The design, however, is not immune to external constraints. In this case the primary constraint is sample size. The project staff initially aimed to enroll 1,800 individuals into the Impact Study over a period of 27 months (October 2015 through December 2017). The realized enrollment rate was more modest than expected. As a result, the Evaluation Team did not have the anticipated statistical power to detect programmatic impacts on the outcomes. Due to the program implementation barriers staff faced in developing referral-producing partnerships with community organizations, staff turnover within the BEAM Team itself, and pushback regarding the randomized study design from AJC partners and eligible candidates, BEAM staff were only able to enroll 365 participants into the ITT (i.e., randomized) sample for analysis. As a result, the Evaluation Team did not have the anticipated statistical power to detect programmatic impacts on the outcomes.

The duration of the study, while not a limitation of the design itself, is a second substantial limitation to fully evaluating the impact of BEAM's GCP intervention on the hypothesized outcomes. A comprehensive evaluation of the impact of BEAM's GCP intervention on educational and employment outcomes would need to be longer than was permitted here. With a maximum of 37 months in the study period (December 2015 through December 2018), participants may not have had enough time to achieve each of the outcomes. The secondary (full-year) analysis on the subsample of participants who would have had a longer window or more time on average in which to achieve outcomes was one attempt to overcome this limitation, but the findings from these analyses are similarly limited by sample size.

The third limitation of this study is related to the data available from the existing administrative data sources (i.e., NSC, NYSDOL). In both educational and employment domains, the Evaluation Team is unable to definitively say whether an individual did not achieve the desired outcome (enroll in postsecondary education, become employed) or whether they did, but outside the reporting jurisdiction of the NSC or NYSDOL. The NSC data used to assess impact on educational outcomes were limited to accredited two- and four-year colleges and universities, and so progress toward non-degree industry certifications was not reflected in the data. Additionally, UI wage data from NYSDOL would not have included wages earned from out of state, self-employment, or federal employment.

To the extent that randomization procedures were followed and attrition is low, the executed Impact Study provides unbiased estimates of the treatment effect. Although a range of rigorous safeguards were instituted to ensure fidelity of randomization, the task was performed on-site by the BEAM OCs, therefore the Evaluation Team cannot be completely sure if all randomization procedures were followed in every case. The aim of this study was to produce empirical, causal responses to the posed research questions, and is just one part of the comprehensive evaluation that the Evaluation Team conducted on BEAM and its GCP intervention. BEAM was a new, innovative program model at the time this evaluation took place. Thus, it is the Evaluation Team's hope that the combination of information gained from the BEAM evaluation will provide BEAM leadership with tools to ask the questions necessary to further develop the program and investigate how it might improve outcomes for postsecondary dropouts.

APPENDIX A: IMPACT EVALUATION METHODS

PURPOSE AND BACKGROUND

The Impact Study aimed to isolate the causal impact that the Bridge to Employment and Academic Marketplace (BEAM) program has on an individual's likelihood of returning to a postsecondary education program, obtaining a degree or certificate, entering employment, and the effect it has on wages within a pre-specified time period. BEAM was designed to provide services that will help youth and adults who have previously attempted and dropped out of college to reenroll in a postsecondary education program and work toward obtaining a degree or certificate that will increase the likelihood of entering employment after program completion. The target population for BEAM was individuals 19 years or older who have previously attempted a postsecondary education training program, but discontinued enrollment prior to completing the program, and who are residents of the nine-county region where BEAM was offered.⁴¹ The Impact Study was a randomized controlled trial (RCT). Outcomes for treatment group members who were offered the defined intervention were compared with those of a control group who were offered American Job Center (AJC) business-as-usual (BAU) program services. Individuals were eligible to participate in the study if they (1) were 19 years or older, (2) resided in one of nine participating counties, (3) had previously attempted college but dropped out before receiving a degree or certificate, (4) gave consent to participate in the study, (5) were willing to return to an AJC to receive services, (6) enrolled to receive services at one of the participating AJCs during the study enrollment period, (7) were not a veteran of the armed forces, and (8) were not currently enrolled in a postsecondary education training program.

The unit of assignment was the individual participant; the unit of analysis was the individual participant. Eligible participants who provided consent during the study enrollment period were randomly assigned to the treatment or control (BAU) condition at a 4:1 ratio (treatment to control). This is an intent-totreat (ITT) study design, which means that the analytic sample consists of all study participants who consented to participate and who were randomly assigned to either the treatment or control condition; as such it does not reflect the entire population of eligible individuals. In an ITT framework, participant outcomes are measured within the treatment condition that they were randomly assigned to, regardless of actual exposure to the assigned intervention. In other words, individuals assigned to receive the Guided Career Pipeline (GCP) program are considered members of the treatment group, regardless of whether they received any GCP services. The Evaluation Team established a data sharing agreement with the New York State Department of Labor (NYSDOL), which is the source of data on participant background characteristics and employment outcomes. Educational enrollment and completion outcome data were obtained from the National Student Clearinghouse (NSC). Data on regional economic conditions were obtained from the U.S. Bureau of Labor Statistics (BLS). All other administrative data were collected and regularly submitted to the Evaluation Team by the OCs. The seven BEAM Outreach Coordinators (OCs) working at the 11 AJCs originally aimed to enroll 1,800 individuals into the Impact Study during 27 months of enrollment (October 2015 through December 2017). Actual enrollment was more modest than expected. The OCs enrolled 365 individuals between December 2015 and March 2018 (27 months).⁴²

 ⁴¹ BEAM was offered in Broome, Chenango, Delaware, Herkimer, Madison, Oneida, Otsego, Tioga, and Tompkins Counties in New York.
 ⁴² Due to lower than expected enrollment, the Evaluation Team and the BEAM Project Director decided to extend enrollment in the Impact Study beyond December 31, 2017 (the original cutoff) to March 31, 2018.

By design, the Impact Evaluation was developed to answer two primary and four exploratory research questions that are concerned with BEAM's effect on outcomes identified by the program's theory of change. As discussed in more detail below, the primary research questions focus on the central aim of the BEAM model – to get academically disconnected adults reenrolled in postsecondary training programs as a means of obtaining industry-recognized credentials that would lead to in-demand, high-paying careers. The exploratory research questions aim to assess program impact on the latter (degree attainment, employment opportunities).

As delineated by the questions below, the Impact Evaluation sought to estimate the impact of BEAM's GCP intervention on participants' likelihood of enrolling in and persisting through a postsecondary education program, obtaining a degree or certificate, entering employment, and receiving higher wages.⁴³ The primary study focused on BEAM's short-term outcomes (postsecondary enrollment, persistence), whereas the exploratory study focused on the program's indirect outcomes (graduation from postsecondary program and increased employment opportunities). Although they serve slightly different purposes, the estimates produced in response to the primary and exploratory research questions are based on the same rigorous methods and therefore are both causal in interpretation.

PRIMARY

- 1. Are treatment group members who are offered GCP more likely to enroll in a postsecondary education program within nine months of enrolling in BEAM than equivalent participants who were offered Career Center Services (CCS)?
- 2. Do treatment group members who are offered GCP remain enrolled in a postsecondary education program for more consecutive semesters than equivalent participants who were offered CCS?

EXPLORATORY

- 3. Are treatment group members who are offered GCP more likely to complete their postsecondary program certificate or degree by the end of the study period than equivalent participants who were offered CCS?
- 4. Are treatment group members who are offered GCP more likely to enter employment within three full quarters of enrolling in BEAM than equivalent participants who were offered CCS?
- 5. Do treatment group members who are offered GCP receive higher wages at the end of the third full quarter post-enrollment than equivalent participants who were offered CCS?
- 6. Do treatment group members who are offered GCP remain employed for more quarters on average than equivalent participants who were offered CCS?

As noted in BEAM's logic model, a large segment of intervention activities is geared toward assisting participants with enrollment in and completion of a postsecondary education program as a means of

⁴³ Note that the Implementation Evaluation section of the full evaluation report discusses the innovative intervention as BEAM, which is the program model that utilizes OCs to provide one-on-one navigation support that is not available through the standard AJC service model. The Evaluation Team does not discuss BEAM as the treatment intervention in the Impact Study section because participants enrolled in the treatment condition could have also benefited from BAU services op top of those offered by BEAM staff. GCP is the package of services that OCs delivered that went above and beyond what would have been available to AJC customers. The GCP package was delivered by BEAM staff and is a core component of BEAM, but is not the whole extent of the BEAM experience. The Impact Study's treatment contrast attempts to isolate the impact of the additional services available to GCP customers that were not available to AJC customers.

increasing subsequent employment opportunities. Time, however, acted as a major constraint in assessing this pathway comprehensively. The total possible duration of time from study enrollment to the latest possible outcome observation was 3 years (36 months). It was expected that participants who engaged in postsecondary education would register for two- or four-year programs. Consequently, a substantial proportion of the sample did not have sufficient time to realize the full extent of the hypothesized education completion and employment outcomes (Research Questions 3 through 6). The ramifications are different for each outcome of interest. Although the Evaluation Team hypothesized that more treatment participants would graduate from postsecondary programs (Research Question 3) than control participants, it was not feasible to accurately measure the full extent of that effect because participants were enrolled in the study for, on average, 22 months. Consequently, the primary research questions focus on the program's effect on a participant's likelihood of enrolling in a postsecondary education program (Research Question 1), and their educational progress, or number of semesters enrolled in the program (Research Question 2). For Research Questions 4 through 6, it was unclear whether time constraints may actually present contradictory findings. That is, because comparatively fewer participants in the control condition were expected to enroll in college, they would have had more time to gain employment, raise their wages, and work a greater number of hours per quarter. As such, Research Questions 3 through 6 are exploratory in nature and are meant to provide additional insight into BEAM's potential effect on desired outcomes.

DATA METHODS

DATA SOURCES AND COLLECTION

Individual-level outcome data, covariate data, and contextual/regional economic data were collected from four main sources for the Impact Evaluation: NYSDOL, the NSC, BLS, and OCs. Participant background characteristics and employment outcome data were provided by NYSDOL; educational outcome data were provided by the NSC. The Evaluation Team used a unique study ID number to merge the participant background and employment data set with the education outcome data set for final analysis.

PARTICIPANT BACKGROUND CHARACTERISTICS AND EMPLOYMENT OUTCOME DATA (DATA SOURCE: NYSDOL) As part of the standard enrollment procedures, all new customers are asked to provide a range of background information requested by NYSDOL's *New Customer Registration Form* and *Supplemental Questionnaire*, which are used universally by all AJCs in New York. Data collected with these forms include date of birth, race and ethnicity, education history and current enrollment status, employment history, veteran status, and disability status. In addition to background information, staff at each AJC are also required to track individual clients' education status, employment history, services provided, and funding sources for those services on an ongoing basis. Data on study participants' background characteristics were collected by the OCs at the time of enrollment in the study (using the *New Customer Registration Form* and *Supplemental Questionnaire*) and entered into the One Stop Operating System (OSOS), a statewide data sharing system used by all AJCs in New York to share customer background, service delivery information, and performance measures with NYSDOL. OCs routinely exported these data to NYSDOL, as is required for all AJC customers.

Prior to the initial implementation of BEAM, the Evaluation Team requested that NYSDOL add four new fields to the OSOS database at each of the 11 participating AJCs. The additional fields capture for each participant enrolled in the study: (1) indicator of participation in the Workforce Innovation Fund (WIF) Round 2 Impact Study, (2) unique study ID number, (3) study group assignment (treatment or control),

and (4) study enrollment date. Capturing this information in OSOS allowed the Evaluation Team to use the unique study ID number to request individual-level background characteristics, services received, and employment outcome data from NYSDOL for all study participants without the need for social security numbers.

Employment outcome data were obtained from NYSDOL in the form of quarterly Unemployment Insurance (UI) wage data. According to NYSDOL, quarterly UI wage records capture 97% of New York State's nonfarm/non-migrant workforce; they do not include self-employed or federal employees. For each participant enrolled in the Impact Study, the evaluators requested four full quarters of UI wage records prior to their date of enrollment and five full quarters of UI wage records after the quarter in which they enrolled in the study. Because enrollment occurred on a rolling schedule, the data time frames vary for each participant.

The Evaluation Team worked with NYSDOL throughout the first and second year of the grant to develop a detailed data sharing agreement. A final, executed agreement was established in November 2016; all necessary precautions to ensure confidentiality and compliance with requirements from the State of New York regarding data security practices were followed.⁴⁴

EDUCATION OUTCOME DATA (DATA SOURCE: NSC)

Education outcome data were requested from NSC. According to its website, NSC houses student outcome data for over 3,600 colleges and universities, accounting for 99% of students enrolled in public and private U.S. institutions.⁴⁵ Coverage is therefore almost – but not entirely – complete. Enrollment verification data include information about the college(s) attended, dates of enrollment, chosen majors, and any degrees earned. The Evaluation Team entered into a data sharing agreement with NSC in January 2016. In order to request individual-level data, the Evaluation Team provided NSC with a list of each enrolled participants' first and last names and date of birth. These data, along with participants' unique study ID number were provided to NSC, which then pulled the requested outcome data and stripped participants' names and dates of birth prior to submitting the data set back to the Evaluation Team. The Evaluation Team used the unique study ID number to merge education and employment outcome data into one data set for analysis.

U.S. BUREAU OF LABOR STATISTICS

In addition to education outcome data collected from NSC and employment outcome data collected from NYSDOL, the Evaluation Team collected publicly available data from the U.S. Department of Labor (USDOL) BLS.⁴⁶ These data incorporated results from the Local Area Unemployment Statistics (LAUS) program and the Quarterly Census of Employment and Wages (QCEW). The Evaluation Team compiled data made available by BLS on the following contextual economic indicators: unemployment rate, total labor force, total employed, and total unemployed (all from the LAUS), and average weekly wages and hours worked (from the QCEW) for each economic quarter between the 3rd quarter of 2015 and the 1st quarter of 2019. The Evaluation Team identified economic conditions for the Metropolitan Statistics

⁴⁴ An amendment to the agreement was executed on August 16, 2018. The amendment revised the list of requested variables and their associated codes to better align with changes made to NYSDOL reporting structure as a result of the implementation of the Workforce Innovation and Opportunity Act (WIOA) in July 2016.

⁴⁵ Retrieved July 16, 2019, from the National Student Clearinghouse, https://studentclearinghouse.org/colleges/studenttracker/

⁴⁶ U.S. Department of Labor, Bureau of Labor Statistics, https://www.bls.gov/data/

Area (MSA) or, if that was not available, for the county in which the AJC was located for the quarter directly preceding the participant's enrollment quarter.

PARTICIPANT ENROLLMENT TRACKING (DATA SOURCE: OUTREACH COORDINATORS) In addition to the information gathered by the *New Customer Registration Form* and *Supplemental Questionnaire*, OCs also recorded the following information for each participant into a Zoho Creator secure webform developed by the Evaluation Team: (1) unique study ID number, (2) study group assignment (treatment or control), (3) study enrollment date, (4) consent date, (5) enrollment site, (6) participant's first and last name, (7) participant's date of birth, and (8) the name of the OC who completed enrollment. This data set served two purposes: (1) it allowed the Evaluation Team to monitor that study enrollment and randomization were being conducted with fidelity, and (2) three of these data elements (unique study ID number, first and last name, and date of birth) were submitted to the NSC to obtain education outcome data for the study sample.

DATA PROCEDURES

As data were obtained in multiple data files, a unique study ID was used to create a composite data set. This ID was created by the Evaluation Team and assigned to each participant by the OC at the time of their enrollment in the Impact Study. Staff at NYSDOL used a secure link to upload encrypted data files to their secure file transfer site, managed by the NYS Department of Taxation and Finance. NYSDOL provided the Evaluation Team with an encryption key to decrypt the data files, which were then stored on a password-protected, limited-access server that requires two-factor authentication for access.

Similar to the employment outcome data, staff at the NSC used a secure link to upload data files to their secure file transfer site and provided the Evaluation Team with unique login credentials to download the data files. Data were then stored on a password-protected, limited access server that requires two-factor authentication for access.

The participant demographic data were submitted in an individual-level, wide format, with one record per individual, identified with the unique ID. The NSC and UI wage data sets were submitted in an individual-level, person-period format, where an individual had multiple observations in the data set based on the number of postsecondary enrollments or quarters of wage records, again identified using the unique ID.

In order to create an analysis-ready data set, the NSC and UI wage data sets were reshaped into a wide format and merged with the demographic data set as well as a data set containing information related to an individual's enrollment in the Impact Study, using the unique ID. The final data set contained, per participant, a single row of data pertaining to an individual's enrollment in the study, educational outcomes, employment outcomes, and demographic data.

VARIABLES

In this section, the Evaluation Team presents a discussion and description of the individual-level covariates, the time-variant economic and contextual variables, and the outcome variables used in the impact analysis.

COVARIATES

Table A1 provides a description of the individual-level covariates that were included in the benchmark analytic models and additional models used to conduct sensitivity analyses. The completeness of covariate data varied across individual variables.

Variable Name	Description of Variable
Gender	Gender is reported as either male or female. A dummy variable was created to report whether an individual was female (1) or not (0). Source: NYSDOL
Age at study enrollment	Age at enrollment is calculated as the length of time between the date the participant enrolled in the study and the participant's date of birth. Source: NYSDOL
White	Dummy variable reporting race as only White (1) or not (0). Source: NYSDOL
Black	Dummy variable reporting race as only Black (1) or not (0). Source: NYSDOL
Multiracial	Dummy variable reporting race as Multiracial (1) or not (0). Source: NYSDOL
Other race	Dummy variable reporting race as another race, not specified by the variables White, Black, or Multiracial (1) or not (0). Source: NYSDOL
Ethnicity	Dummy variable reporting ethnicity as Hispanic (1) or not (0). Race and Ethnicity were reported separately. Source: NYSDOL
Disability status	Disability status is reported as having any disability, as defined in Section 3(2)(a) of the Americans with Disabilities Act of 1990 (42 U.S.C. 12102). Recognized disabilities include a physical or mental impairment that substantially limits one or more of the person's major life activities.
	Source: NYSDOL
Highest level of education completed	Highest level of education is reported as the highest degree attained (high school or less, some college but no degree, or vocational/technical certificate, Associate's Degree or higher) prior to enrolling in the study.
	 Three dummy variables were created to report whether an individual's highest education level was: High school diploma or less (1), or not (0) Some college (1), or not (0) College degree (1), or not (0) Source: NYSDOL
Low income	As defined by NYSDOL, participants are considered to have low income if any of the following apply: (1) participant or family has recently received (in the past 6 months) or is receiving SNAP, TANF, SSI, or state or local public assistance; (2) is in a family with total family income that does not exceed the higher of the poverty line or 70% of the lower living standard income level; (3) is a youth who receives or is eligible to receive a free or reduced price lunch; (4) is a foster child on behalf of whom state or local government payments are made; (5) is a participant with a disability whose personal income is below the poverty line (i.e., family income is not taken into account); (6) is homeless; (7) is a youth living in a high-poverty area.
	A dummy variable was created reporting an individual as meeting any of the above criteria (1) or not (0). Source: NYSDOL
Random block	Randomization block is reported as the AJC site at which the participant was enrolled in the study or the OC who enrolled the participant in the study.
	A series of dummy variables were created to report whether an individual was enrolled at a particular site (1) or not (0). Source: Enrollment Log
Average quarterly wages prior to enrollment	The mean of the certified quarterly wages earned for the four quarters immediately prior to enrollment in the study. Source: NYSDOL

Table A1. Covariate Variables

BEAM IMPACT EVALUATION

Variable Name	Description of Variable
Average wages during each quarter prior to enrollment	Certified quarterly wages earned for each of the four quarters prior to enrollment in the study. Source: NYSDOL
Average employment status prior to enrollment	The proportion of quarters that a participant was employed out of the four quarters prior to study enrollment. Participants are considered to be employed in a particular quarter if they earned any wages (more than \$0) during that quarter. Source: NYSDOL
Employment status during each quarter prior to entry	A series of four dummy variables indicating if the individual was employed (1) or not (0) in a particular quarter; a dummy is constructed for each of the four quarters prior to enrollment in the study. Source: Wage data set
Homeless	Dummy variable indicating if the individual is homeless at entry (1) or not (0). Source: NYSDOL
Offender	Dummy variable indicating if the individual is an offender at entry (1) or not (0). Source: NYSDOL
SSI/SSDI recipient	Dummy variable indicating if the individual has received either SSI or SSDI ⁴⁷ benefits at entry (1) or not (0). Source: NYSDOL
TANF recipient	Dummy variable indicating if the individual has received TANF benefits at entry (1) or not (0). Source: NYSDOL
UI status	Dummy variable indicating if the individual has received UI benefits at entry (1) or not (0). Source: NYSDOL
Funding stream	A series of three dummy variables indicating whether an enrollment is a WIOA ⁴⁸ adult (1) or not (0), a WIOA dislocated worker (1) or not (0), or a WIOA youth (1) or not (0). Source: NYSDOL

ECONOMIC AND CONTEXTUAL INDICATORS

Economic and contextual variables captured second-order processes existing outside of the control of the study design that may have influenced outcomes for participants who were entering the study at different times and in different locations. Specifically, economic and contextual variables attempted to capture differences in labor market conditions at different points in time during the study period. Including these variables can help diminish any potential bias stemming from variable economic conditions across time and geographic location. **These variables were not used in the benchmark confirmatory or exploratory analytic models; they were used to conduct a sensitivity analysis for each outcome.** The time-variable and contextual indicators are detailed in Table A2.

⁴⁷ Social Security Disability Insurance.

⁴⁸ Workforce Innovation and Opportunity Act.

Table A2. Economic and Contextual Variables

Variable Name	Description of Variable
MSA/County unemployment rate for quarter preceding enrollment	A continuous variable describing the unemployment rate for the MSA/county where the participant resided for the quarter preceding the participant's enrollment in the study. Source: BLS; NYSDOL
MSA/County labor force for quarter preceding enrollment	A continuous variable describing the labor force for the MSA/county where the participant resided for the quarter preceding the participant's enrollment in the study. The labor force is the sum of employed and unemployed persons residing in the area. Source: BLS; NYSDOL
MSA/County number employed for quarter preceding enrollment	A continuous variable describing the number of employed persons for the MSA/county where the participant resided for the quarter preceding the participant's enrollment in the study. Source: BLS; NYSDOL
Seasonal quarter of enrollment	A series of four dummy variables describing the seasonal quarter that the participant enrolled in the study; variables constructed as follows: - Enrolled during Jan–Mar (1), or not (0) - Enrolled during Apr–Jun (1), or not (0) - Enrolled during Jul–Sep (1), or not (0) - Enrolled during Oct–Dec (1), or not (0) Source: Enrollment Log
Month of study entry	A continuous variable describing the number of months that had elapsed in the study period at the point of the participant's enrollment in the study, where 1 = the first month of the study window (Dec 2015) and 37 = the last month of the study window (Dec 2018). Source: Enrollment Log
Days into quarter of enrollment	A continuous variable describing the number of days that have elapsed since the start of the quarter that the participant enrolled in the study to the date they enrolled in the study. Source: Enrollment Log
Average hourly earnings for quarter preceding enrollment	A continuous variable describing the average hourly earnings for the state of New York for the quarter preceding the participant's enrollment in the study. Source: BLS
Average weekly hours for quarter preceding enrollment	A continuous variable describing the average number of hours worked each week for the state of New York for the quarter preceding the participant's enrollment in the study. Source: BLS
State minimum wage for quarter preceding enrollment	A continuous variable describing the statewide minimum wage in New York during the quarter preceding the participant's enrollment in the study. Source: BLS

OUTCOME VARIABLES

Educational outcomes were operationalized with three measures: (1) enrollment into a postsecondary education program, (2) number of consecutive semesters enrolled, and (3) graduation from postsecondary education program. Employment outcomes were assessed with three measures: (1) employment status (i.e., whether one was employed) during any of the first three full quarters after enrollment; (2) wages (i.e., the total wages earned) in the third full quarter after enrollment; and (3) quarters employed (i.e., total number of quarters that an individual was employed) during the first three full quarters after enrollment. Table A3 outlines how these outcome measures were constructed.

Table A3. Time-Vari	ant Economic a	and Contextual	Variables
---------------------	----------------	----------------	-----------

Outcome Name	Description of Outcome	Timing of Measure						
Enrollment in postsecondary education program	stsecondary education The outcome was measured as the probability of enrolling in a postsecondary education institution within 9 months of study enrollment, as reported by the NSC. Measured at the month follow-u (between days the date the pa operalled in the							
	 The outcome variable was constructed using the following data elements in the NSC data set: Record found (Y/N) Enrollment begin date 							
	And the following data element in the Enrollment Log data set: • Study enrollment date							
	A participant was considered enrolled in postsecondary education program within 9 months of study enrollment if he/she had a postsecondary education enrollment begin date that was both after their study enrollment date and before the end of their nine-month follow-up window.							
	The measure is operationalized as a dummy variable indicating if the participant enrolled in a postsecondary education program (1) or not (0).							
Consecutive semesters enrolled in postsecondary education	The outcome was measured as the number of semesters (spring, summer, fall) that a participant was continuously enrolled in any postsecondary training program during the study period, as reported by the NSC.	Measured at the end of the study period (December 31, 2018)						
	 The outcome variable was constructed using the following data elements in the NSC data set: Enrollment begin date Enrollment end date 							
	And the following data element in the Enrollment Log data set: • Study enrollment date							
	The measure is operationalized as a count variable of the number of semesters a participant was continuously enrolled in any postsecondary education program ranging from 0 (never enrolled) to 7 (enrolled for seven consecutive semesters). A participant was considered continuously enrolled if (1) the data indicated he/she was enrolled for two out of the three semesters in each calendar year they were enrolled in the postsecondary program and (2) they did not skip two semesters in a row.							
Graduation from postsecondary education program	This outcome was measured as the probability of graduating (i.e., earning a recognized degree) from a postsecondary education program by the end of the study period, as reported by the NSC.	Measured at the end of the study period (December 31, 2018)						
	 The outcome variable was constructed using the following data elements in the NSC data set: Graduated Graduation date 							
	The measure is operationalized as a dummy variable indicating if the participant graduated from a postsecondary education program (1) or not (0).							

Outcome Name	Description of Outcome	Timing of Measure
Employed in any of three quarters after enrollment	This outcome was measured as the probability of being employed for at least one quarter during any of the three full quarters after study enrollment. Employment was indicated if the individual had any amount of wages reported in the quarters being measured.	Measured during the 1 st , 2 nd , and 3 rd full quarter after study enrollment
	The outcome variable was constructed using quarterly UI wage data from NYSDOL.	
	The measure is operationalized as a dummy variable indicating if the participant was employed for at least 1 quarter after study enrollment (1) or not (0).	
Wages during the 3 rd full quarter after enrollment	This outcome was measured as the total wages reported for the 3 rd full quarter following study enrollment.	Measured during the 3 rd full quarter after the participant enrolled in the study
	The outcome variable was constructed using quarterly UI wage data from NYSDOL.	
	The measure is operationalized as a continuous variable that represents the total wages reported for the 3 rd full quarter after study enrollment. Wages reported from multiple employers were summed to create a total for the quarter.	
Number of quarters employed	This outcome was measured as the total number of quarters (range 0 to 3) that the participant was employed after he/she enrolled in the study. Employment was indicated if the individual had any amount of wages in the quarters being measured.	Measured during the 1 st , 2 nd , and 3 rd full quarters after study enrollment
	The outcome variable was constructed using quarterly UI wage data from NYSDOL.	
	The resulting variable is a count variable of the number of quarters that the individual was employed during the 1^{st} , 2^{nd} , and 3^{rd} full quarters after study enrollment. A value of 0 means the individual was unemployed in all 3 quarters, and a value of 3 means the individual was employed in all 3 quarters.	

MISSING DATA

The Evaluation Team did not collect data from study participants directly; all outcome data were derived solely from existing administrative data sources (NYSDOL UI wage records, NSC). In general, the completeness of these data is known. According to data available on the NSC website that breaks down enrollment coverage by state, the NSC database covered enrollment at 97.1% of accredited postsecondary institutions in New York on average between fall 2015 and fall 2018, with more complete data for public institutions compared with private, for-profit institutions.⁴⁹ NYSDOL UI wage records cover 97% of non-farm/non-migrant workers in the state, with self-employed and federal government employees composing the largest segment of missing data.⁵⁰ In addition, these data do not include residents of the state who work out-of-state. In both sets of data, however, nonexistence of data and non-incidence of the outcome (enrollment, employment) are not differentiated. The study participant who fails to enroll or become employed was coded the same as one who does, but outside the scope of the data collection efforts. So, while the general completeness of the data is known, the Evaluation Team is not able to calculate overall or differential missing data figures specifically. Outcome estimates

⁴⁹ Data retrieved July 16, 2019, from https://nscresearchcenter.org/workingwithourdata/

⁵⁰ Information retrieved July 16, 2019, from https://labor.ny.gov/data-sharing/data-sharing-faqs.shtm

may be lower than they actually appear, but owing to the randomization of participants, the Evaluation Team assumes the influence on treatment and control outcomes is equal. As such, the difference between the two should not be biased because both groups should experience an equal amount of missing data. Incomplete covariate data were included in analytic models and missing cases were imputed to the grand mean.

METHODS

APPROACH

The purpose of the Impact Evaluation was to assess whether offering BEAM's GCP intervention to postsecondary dropouts improved the likelihood that those individuals would reenroll in postsecondary training, persist longer, graduate from postsecondary training, become employed, earn higher wages, and remain employed for more quarters. These questions are investigated using an RCT design. As specified in the Evaluation Design Report (EDR), the Evaluation Team was interested in estimating the impact of BEAM across all four workforce regions that implemented the program.

STUDY DESIGN

The Evaluation Team designed an RCT to provide a causal response to the research questions outlined above. From the standpoint of internal validity, causal inferences are justified in a well-executed RCT, because random assignment ensures that the treatment and control groups are equal in expectation in terms of observable and unobservable characteristics, as well as any environmental factors. The key assumption in this expectation is that the analytic sample (the sample of individuals included in the analysis) is sufficiently representative of the ITT sample (sample of those randomized).

The Impact Study contrasts the employment and educational outcomes of individuals randomly assigned to the treatment condition (GCP) with those of individuals assigned to receive services that they would typically be offered by AJC (CCS control group). The purpose of the study is to determine the program's efficacy in improving these outcomes for AJC clients. Participants were randomly assigned to treatment or control conditions at a 4:1 ratio (where 289 were enrolled in treatment and 76 were enrolled in control). Eligible participants were assigned to treatment or control condition by a random-order scheme implemented by the OCs at each of the 11 participating AJCs.

CONTROL EXPERIENCE - BUSINESS AS USUAL

In contrast with those participants who were offered GCP by OCs, participants assigned to the control condition were provided education and employment assistance from the AJC staff that is identical to what they would have received at each of the AJCs prior to the advent of BEAM. This is a business as usual (BAU) contrast. Specific services that are typically provided include job readiness assessments, job counseling, skill assessment, career planning, referrals to and funding for training, computer and Internet access for resume development and job searching, printing services, and library access. The BAU experience of the control group approximates the counterfactual of interest as closely as possible. The BAU experience would have naturally varied by AJC, county, Workforce Development Board (WDB) region, and participants' educational and career goals – there were 11 AJCs, across 9 counties, within 4 WDBs, and participants entered the study at varying degrees of financial stability – but participants

assigned to the control condition would have received the same core structure of services under AJC funding streams (WIOA, the Wagner-Peyser Act, etc.)⁵¹

Using a BAU contrast rather than a no-intervention control is useful because it contrasts the experimental intervention (in this case GCP, as delivered by OCs) with current practice. This means the outcomes that result from the BAU experience represent the expected outcomes for the current AJC experience. The impact estimates produced by this study, therefore, are best conceived as the average "improvement" in outcomes that are attributable to BEAM above and beyond what is being achieved by current practice on average. In other words, a null outcome does not represent lack of improvement in participant outcomes, but rather no relative improvement over what is currently being achieved by BAU at AJCs. A positive impact represents a desirable improvement over current practice and a negative impact represents an undesirable decline over current practice.

ELIGIBILITY AND EXCLUSION CRITERIA

The Impact Study aimed to make inferences about program impacts on clients who were offered the opportunity to enroll in GCP and receive individualized assistance from OCs. Because participation in the study was voluntary, the inference space does not extend to the entire AJC customer population, but rather a sample of those who were eligible and consented to participate. To be enrolled in the study, individuals had to either appear at an AJC that was offering the intervention during the enrollment period or meet in-person with an OC at a community-based location, meet eligibility criteria, and consent to participate. Any individual who met the eligibility criteria was asked if he/she would like to hear more about the study. If they said yes, the OCs discussed the consent procedures to allow the individual to make an informed decision about study participation. If the individual consented, they were enrolled in the study, randomized into an intervention condition (treatment or control), and considered part of the ITT sample. Note that once an individual was enrolled in either condition, he/she was considered a study participant in that condition regardless of the actual exposure to those conditions.

A number of criteria were established for participation in the study. To be eligible to enroll in the study, customers had to:

- 1. Be age 19 or over⁵²
- 2. Reside in one of the following nine counties in New York: Broome, Chenango, Delaware, Herkimer, Madison, Oneida, Otsego, Tioga, Tompkins
- 3. Have previously attempted college but dropped out before completing a program of study⁵³
- 4. Be determined by an OC to be a good fit for the study (i.e., likely to return to AJC to receive services)
- 5. Enroll to receive services at one of the participating AJCs during the study enrollment period

⁵¹ For the sake of simplicity, any mention of WIOA programming or funding streams with regard to the Impact Study implies the WIA equivalent when considering data prior to the implementation of WIOA in July 2016.

⁵² The original eligibility specified that individuals had to be between 19 and 24 years old to be eligible for the study. Grant modification 1 increased the age limit to 30 years old in August 2015. Grant modification 5 (November 2016) lifted the age cap entirely such that any individual over 19 years old would meet that criterion for the study.

⁵³ An individual could have previously completed a postsecondary program and received a degree or certificate for that program, enrolled in a higher-level program and subsequently dropped out before receiving the respective degree or certificate. For example, an individual may have received an Associate's Degree, but dropped out before completing a Bachelor's-level program.

- 6. Not be a veteran of the armed forces⁵⁴
- 7. Not be currently enrolled in a postsecondary education program
- 8. Give consent to participate in the study

ASSIGNMENT PROCEDURES

The Impact Study was an RCT where the unit of assignment and the unit of analysis were the individual participant. If a participant met all eligibility criteria, the OC would assign him/her the next available study ID number and use a sealed "assignment envelope" to enroll the participant into either the treatment or control group. The envelopes were prepared by the Evaluation Team and were sealed until each study enrollment meeting. Prior to implementation, and on an ongoing basis as needed, the Evaluation Team prepared a batch of random assignment envelopes for each of the seven BEAM OCs.⁵⁵ Each envelope contained a piece of paper indicating the experimental condition (GCP or CCS) sealed inside. Study IDs were demarcated on the outside of the randomization envelopes, and were placed in order from smallest to largest, such that the smallest ID was to be assigned first and the largest ID was to be assigned last. The Evaluation Team kept a record of the original randomization sequence (the order of the IDs and the condition to which each ID was assigned) to ensure that procedures were adhered to and randomization was carried out with fidelity. The number of treatment and control condition envelopes prepared for each site adhered to the assignment ratio of 4:1 (treatment to control). The Evaluation Team used a random number generator in Stata 14 to create a randomly ordered string of Ts and Cs at a 4:1 ratio. The Evaluation Team randomly assigned blocks of participants, such that for every five participants there were exactly four treatment assignments and one control assignment in that block. To ensure that OCs were not able to anticipate the final assignment in a block, block sizes were alternated such that some assigned 15 participants in random-order blocks (12 treatment and 3 comparison assignments), others assigned 10 participants in a block (8 treatment and 2 comparison), and others assigned 20 participants in a block (16 treatment and 4 comparison). Each assignment allocation was then placed in the numbered envelopes in the original randomly assigned order. Before the envelopes were sealed, the Evaluation Team recorded in the Randomization Log the assigned study ID number and the random treatment allocation to use for reference when checking that study IDs and experimental conditions were being assigned to new participants in the intended order. The Evaluation Team regularly (e.g., monthly) verified that new participants were enrolled in the study with fidelity by cross-referencing the Enrollment Log maintained by OCs in Zoho Creator against the Randomization Log.⁵⁶

⁵⁴ Any individual who is a veteran of the armed forces was not eligible to participate in the study, as determined by the USDOL in the WIF Round 2 solicitation of grant applications. Any such individual was referred to needed program services, including those in both the treatment and control condition but was excluded from the Impact Study.

⁵⁵ At the start of the Impact Study, OCs had to screen and enroll potential participants on-site at one of the 11 AJCs participating in BEAM. During this time, the Evaluation Team created a batch of randomization envelopes for each of the 11 AJCs, that the OC would store and use at that site. Grant modification 4 (May 2016) revised the study protocols such that OCs could screen and enroll potential participants off-site at community locations (e.g., educational institution, library, community-based organization, etc.). After this modification, the Evaluation Team instructed the OCs to destroy (e.g., shred) the remaining randomization envelopes at each site, and provided each of the seven OCs with a new batch of envelopes that they could keep with them as they traveled to different sites.

⁵⁶ Although the Evaluation Team set up safeguards to regularly monitor that randomization was being conducted with fidelity, there were limitations to this effort that the BEAM staff could have potentially exploited. For example, a scheming OC could have opened up all of their randomization envelopes ahead of time to ensure that ideal candidates for GCP were selected for enrollment at the appropriate time. The Evaluation Team assessed the Impact Study sample for baseline balance on key characteristics and, for the most part, did not find evidence of tampering. In addition, the Evaluation Team regularly provided study procedure trainings to new and established OCs throughout the study period to reinforce fidelity.

Blocking occurred in two phases during the span of the Impact Study. Between December 2015 and July 2016, blocking occurred at the AJC site where the individual was enrolled in the study. The random sequence of treatment and control allocations were generated for each of the 11 sites because customers were required to meet in-person with an OC at an AJC in order to be screened for and enroll in the study. Due to low enrollment rates at the sites, the BEAM Project Director requested grant modification 4 in May 2016, which would allow new study enrollments to be conducted at community-based sites, such as postsecondary educational institutions, community-based organizations, job fairs, libraries, etc. This change was made in an effort to increase the rate at which OCs were enrolling new participants by allowing them to reach a larger pool of potentially eligible individuals who may not have shown up at an AJC to receive services prior to their enrollment in the study. As a result, the second phase of random assignment blocking was done at the level of the OC conducting the enrollment. The Evaluation Team created new batches of assignment envelopes for each of the seven OCs. In the event that BEAM experienced staff turnover between July 2016 and the end of the enrollment period in March 2018, the remaining envelopes were used by the new OC who was hired to replace the open position.

At the time of enrollment, the OC would open the envelope that was labeled with the next study ID number in the sequence; the study ID number written on the outside of the envelope was assigned to the individual study participant and the paper inside indicated the condition to which the participant was assigned. The OCs were trained to always use the assignment envelope with the lowest number so that participants were enrolled in the study with consecutive, ascending study ID numbers. The OCs would then enter enrollment and assignment data into the *Enrollment Log*.

ENROLLMENT PROCEDURES

Prior to screening and enrolling an individual in the study, the OCs were instructed to first check the OSOS database (either at the AJC or on their laptops if off-site) to determine that the individual had not already been enrolled in the study by another OC. If the individual was a new customer, the OC checked that he/she met all of the eligibility requirements mentioned above. If not eligible, the OC referred the individual to needed services at the AJC.⁵⁷ If eligible, the OC briefly explained the study and asked if the person was interested in participating. If yes, the OC obtained written, informed consent from the participant.

As noted above, enrollment in the study occurred during in-person appointments between the potential participant and the OC, either at one of the 11 AJCs or at an external community-based location. The enrollment procedures occurred in the following order: (1) OC completes eligibility screening; (2) OC obtains written, informed consent from customer; (3) OC assigns eligible participant a unique study ID number and randomly assigns participant to the treatment or control group using assignment envelope; (4) participant completes the NYSDOL *Customer Registration Form* and *Supplemental Questionnaire* (this is true for participants in both conditions); (5) OC enters participant enrollment data into the *Enrollment Log*; and (6) OC enters participant data from the *Customer Registration Form*, *Supplemental Question*, and study enrollment data into OSOS.

⁵⁷ In almost all cases, customers who did not consent to participate in the Impact Study were not eligible to receive GCP services. The only exception to this rule was in the case of veterans. Grant requirements dictated that veterans be excluded from random assignment procedures and offered the intervention, if they desired. For this reason, the Evaluation Team excluded veterans from participation in the Impact Study.

ANALYSIS PLAN

The Impact Study investigates the extent to which offering the treatment intervention to participants has a causal effect on their likelihood of enrolling in and completing a postsecondary education program by obtaining a degree or certificate, their likelihood of entering employment, and increasing their average wages. The analysis is completed within the ITT framework, which does not measure the effect of the participant's exposure to the treatment itself, but rather the effect of the offer of the treatment relative to the offer of receiving the control condition (BAU). That is to say, the causal effect that is operationalized is the actual policy option - the offer to receive one of two intervention conditions (treatment/control). Under this structure, the Evaluation Team is able to produce an unbiased estimate of the treatment effect regardless of participant crossover and variation in program exposure, with the assumption that randomization was done with fidelity and attrition remains low. To validate the fidelity of randomization – in terms of the measured variables – the Evaluation Team produced diagnostics of baseline equivalence of the analytic samples. Estimates of program impact on each of the six outcomes were produced by way of regression equations that model outcomes as a function of treatment assignment, blocking variables, and covariates. Regression modeling is not necessary to reduce bias; however, it is useful to increase the precision of impact estimates. To address the question of meaningfulness of estimate, impact estimates were transposed into standardized effect sizes to contextualize the results.

ASSESSMENT OF BASELINE EQUIVALENCE

The Evaluation Team assessed baseline equivalence of treatment and control groups within the analytic sample by assessing the pre-intervention differences in important background characteristics and educational and employment outcomes observed in the data. To assess equivalence, the Evaluation Team generated a model-based estimate of the baseline difference between treatment and control groups on pre-intervention variables; the empirical model is a reduced form of the model used to estimate program impact (as specified in the Analytic Model Specifications section below). It is a reduced form because individual-level covariates are omitted. Separate models were run, and estimates provided, for each of the variables selected for baseline equivalence testing within the analytic sample. Where the baseline variable is continuous, the model is estimated with OLS and the standardized difference is calculated using the Hedges' *g* formula; where the baseline variable is dichotomous, the model is estimated with a multilevel logistic model and the difference in the probability of the occurrence is calculated with the Cox Index formula.^{58, 59}

⁵⁸ The EDR proposed to estimate impacts of dichotomous outcomes with the linear probability model (LPM). The advantages of the LPM include ease of interpretation and computational flexibility (OLS does not drop observations when a fixed-effect block indicator predicts the outcome perfectly). If probabilities are moderate (.20 < *p* < .80), the LPM and OLS fit the data equally well, violations of OLS assumptions tend to be minor, and both will produce substantively comparable estimates. In this case the LPM should be preferred for its ease and flexibility. When probabilities range higher (>.80) or lower (<.20), the logistic model will be preferred to OLS. For many of the baseline covariate variables probabilities are moderate (e.g., prior employment, gender, race/White indicator, low income status), but for others (e.g., race/other, race/Multiracial, SSI/SSDI recipient, TANF recipient) the probabilities are much lower (<.20). Consequently, the Evaluation Team made the decision to estimate baseline equivalence statistics for dichotomous baseline variables (standardized mean difference – in terms of a Cox Index score) using a logistic regression model.

⁵⁹ The EDR proposed fixed effects to account for blocking by the AJC site. Because the assignment ratio was equal at all sites at all times, this statistical adjustment was intended as a means to improve the precision of model estimates. The initial plan was premised on the expectation that 1,800 participants would be enrolled by the program across 11 sites. Enrollment was more modest than expected, however, and as such some sites enrolled a very small number of participants. Kahan (2014) recommends adjusting for site effects with random rather than fixed effects when the number of sites is relatively large compared with the overall sample size. First, a fixed-effects model assumes that the overall sample size is large and that the sample size compared to the number of centers is large. As a result, it requires the exclusion of centers where all participants either experience the same outcome or are members of the same treatment group, reducing the overall sample size. Second, when this assumption is not met and there is a small number of participants per site, fixed-effects models can produce biased estimates of

CONTINUOUS VARIABLES

The following model was used to produce estimates of baseline equivalence for continuous variables:

$$Y_{baseline} = \beta_0 + \beta_1 T + \sum (\beta_P Block_P) + \varepsilon$$

Where:

*Y*_{baseline} is the baseline measure of the diagnostic variable being used to establish baseline equivalence;

T is a dummy treatment indicator variable whose value equals 1 if the participant is randomized into the treatment group and zero otherwise;

Block is an *n*–1vector of blocking dummy variables that are coded 1 if the participant was enrolled at site/by OC *n* and coded 0 otherwise;

 β_0 – the intercept term, which represents the adjusted mean value of the baseline measure for participants in the control sample, with all other variables in the model held constant at zero;

 β_1 – this represents the adjusted (but not standardized) mean difference in the baseline equivalency variable between treatment and control participants.

The standardized differences of means are reported for each diagnostic variable for the analytic sample for the treatment and control groups. For diagnostic variables that are continuous measures, Hedges' *g* is used to calculate the standardized difference of means. The calculation of the standardized mean difference is as follows:

$$g = \frac{\beta_1}{S_p}$$

Where β_1 is the adjusted mean difference in the variable used to establish baseline for the treatment and comparison groups (calculated in the first step) and S_p is the pooled standard deviation.

DICHOTOMOUS VARIABLES

The following model is used to produce estimates of baseline equivalence for dichotomous variables:

Level 1: Participants

$$logit(Y_{baselinejk}) = \beta_{0k} + \gamma_{1k}(T_{jk}) + \varepsilon_{jk}$$

Level 2: Randomization Block

$$\beta_{0k} = \gamma_{00} + \gamma_{02}(Block_k) + \mu_{0k}$$

treatment effect or inflated type I error rates. As a result, Kahan recommends the use of a random effects model under these circumstances, which is able to retain all sites (and as such, all individual participants) in the analysis, even when they have a small, homogenous sample, and because it is likely to produce a valid estimate of the treatment effect, even when the site effects are not normally distributed. Though it is desirable to have a baseline equivalence model mirror an impact estimate model, the Evaluation Team determined that it is preferable to use a multilevel logistic model to estimate baseline equivalence balance statistics, while maintaining the use of a fixed-effects LPM to estimate program impact. See Kahan, B. C. (2014) Accounting for center-effects in multicenter trials with a binary outcome – when, why, and how? *BMC Medical Research Methodology*, 14:20.

Where:

 $Y_{baseline}$ is the measure of the diagnostic variable being used to establish baseline equivalence for participant *j* at site *k*;

 γ_{00} is the estimated mean score of the baseline equivalency measure for all control participants, controlling for other variables in the model;

 γ_{1k} is the adjusted (but not standardized) mean difference in the baseline equivalency variable between treatment and control participants; this is the estimate of interest for the assessment of baseline equivalence;

 T_k is a dummy intervention indicator variable for participant *j* at site *k* whose value equals 1 if the participant is randomized into the intervention group and 0 otherwise;

 $Block_k$ is a set of n-1 dummy variables to capture the variable effects of the n sites at which the participants were enrolled;

 μ_{0k} is the un-modeled site-level variability that is unexplained by level-2 predictors.

For dichotomous variables, the difference in the probability of the occurrence of an event is calculated and then the Cox Index is used to standardize that measure.

For dichotomous measures, the difference in group means is calculated as the difference in the probability of the occurrence of an event. The effect size measure of choice for dichotomous outcomes is the Cox Index, which yields effect size values similar to the values of Hedges' g that one would obtain if group means, standard deviations, and sample sizes were available, assuming the dichotomous outcome measure is based on an underlying normal distribution. Defining p_t and p_c as the probability of an outcome for participants in the treatment and control groups, the effect size is given by:

$$d_{cox} = \left[ln\left(\frac{p_t}{1-p_t}\right) - ln\left(\frac{p_c}{1-p_c}\right) \right] / 1.65$$

ANALYTIC MODEL SPECIFICATIONS

The Evaluation Team estimated program impacts using a regression equation that modeled outcomes as a function of treatment status, the baseline measure of the outcome variable (or a proxy), and other covariates. While a straight difference-of-means/proportion approach would have provided unbiased estimates of the effect of the treatment intervention, a model-based approach is preferred with covariates because it increases the precision of those estimates. An OLS model was used to estimate all outcomes (using Stata 15).⁶⁰

The following model was used to estimate the impact of BEAM on all outcomes:

⁶⁰ The EDR proposed to estimate impacts of dichotomous outcomes with the LPM. The advantages of the LPM include ease of interpretation and computational flexibility (OLS does not drop observations when a fixed-effect block indicator predicts the outcome perfectly). As such, the benchmark approach reported in the Impact Study findings uses OLS/LPM for all outcomes. The Evaluation Team conducted a series of sensitivity analyses to test for variations and/or violations in statistical assumptions. For dichotomous outcomes, additional analyses were conducted using logistic regression; whereas for count outcomes, an array of count models were run. See Appendix C for additional details of these, and other, sensitivity analyses conducted.

$$Y_{Post} = \beta_0 + \beta_1 T + \beta_2 Y_{Pre} + \beta_{n-1} Block_{n-1} + \sum (\beta_P X_P) + \varepsilon$$

Where:

*Y*_{Post} is the outcome variable;

 Y_{Pre} is the baseline measure of the outcome variable;

T is a dummy treatment indicator variable whose value equals 1 if the participant is randomized into the treatment group and 0 otherwise;

Block is an *n*–1 vector of indicator dummy variables that are coded 1 if the participant was enrolled at site/by OC *n* and coded 0 otherwise, and centered at the grand mean;

 X_P is a *p* vector of baseline (i.e., measured prior to receiving intervention or exogenous to treatment) participant-level covariates to account for the variation in outcomes associated with these groups;

 β_0 is the constant term, which represents the baseline regression-adjusted mean of the outcome variable for the comparison group, with all other variables in the model held constant at zero;

 β_1 is the parameter estimate of substantive interest and represents the mean difference in the outcome of interest for those in the treatment condition.

Statistical significance is based on test statistics produced by Stata 15 for the coefficient β_1 using a two-tailed test, with p < .05.

ANALYTIC SAMPLES

BASELINE EQUIVALENCE

This section presents a description of the demographic characteristics of study participants and the baseline balance statistics for the treatment and control groups in the form of standardized mean differences (continuous) and differences in probability of occurrence (dichotomous). Results are provided separately for the benchmark analytic sample (full ITT sample) and the subsample of participants who were enrolled in the study by December 2017 and who are included in secondary analyses. Researchers are encouraged to assess baseline equivalence with standardized difference statistics rather than hypothesis tests, such as a t-test. Although there is no consensus on what value denotes balance, the *What Works Clearinghouse* specifies that differences less than or equal to 0.05 standard deviations require no statistical adjustment for groups to be considered equivalent.⁶¹ For differences between 0.05 and 0.25 standard deviations, an analysis must include an acceptable statistical adjustment for the baseline characteristic to meet equivalence standards. Differences above 0.25 standard deviations in value indicate nonequivalence of groups on that baseline characteristic.

Table A4 presents the treatment and control group means for each characteristic and the balance statistic in the form of standardized differences for the benchmark analytic sample. Figure A1 presents the standardized differences graphically. The vertical grey dashed lines within the figure indicate

⁶¹ What Works Clearinghouse Standards Handbook Version 4.0. Retrieved July 16, 2019, from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_standards_handbook_v4.pdf

standardized differences that are equal to or less than 0.1 and 0.25 standard deviations, which represent a region of acceptable balance in applied research.⁶²

Characteristic	Treatment	Control	Standardized Difference
Age	(n = 289)	(<i>n</i> = 76)	
Mean age in years at enrollment	30.2	32.0	-0.21
Gender	(<i>n</i> = 268)	(n = 75)	
Female	62.7%	69.3%	-0.18
Race	(n = 253)	(n = 73)	
White	65.2%	72.6%	-0.24
Black	26.9%	21.9%	0.18
Multiracial	6.3%	4.1%	0.28
Other	1.6%	1.4%	0.07
Ethnicity	(<i>n</i> = 222)	(<i>n</i> = 62)	
Hispanic/Latino	10.4%	9.7%	0.04
Low income status	(<i>n</i> = 289)	(<i>n</i> = 76)	
Yes	54.0%	53.9%	-0.02
Receives SSI or SSDI	(<i>n</i> = 289)	(<i>n</i> = 76)	
Yes	6.2%	6.6%	-0.04
Receives TANF	(<i>n</i> = 289)	(<i>n</i> = 76)	
Yes	10.7%	11.8%	-0.10
Individual with a disability	(<i>n</i> = 258)	(<i>n</i> = 69)	
Yes	31.4%	15.9%	0.53
WIOA recipient	(<i>n</i> = 289)	(<i>n</i> = 76)	
Adult	76.1%	77.6%	-0.06
Youth	8.3%	3.9%	0.48
Dislocated worker	27.3%	23.7%	0.12
Highest educational level	(<i>n</i> = 289)	(<i>n</i> = 76)	
HS Diploma or equivalent	35.6%	38.2%	-0.08
Some college	51.9%	44.7%	0.18
Associate's Degree or higher	12.5%	17.1%	-0.22
Proportion employed pre-enrollment	(<i>n</i> = 289)	(<i>n</i> = 76)	
1 st quarter	60.6%	53.9%	0.16
2 nd quarter	62.3%	55.3%	0.18
3 rd quarter	60.6%	50.0%	0.26
4 th quarter	60.2%	52.6%	0.19
Mean quarterly wages pre-enrollment	(<i>n</i> = 289)	(<i>n</i> = 76)	
1 st quarter	\$2,590.45	\$2,272.62	0.10
2 nd quarter	\$2,787.09	\$2,531.38	0.08
3 rd quarter	\$2,939.42	\$2,167.78	0.22
4 th quarter	\$3,061.62	\$2,279.99	0.21
Regional labor statistics	(<i>n</i> = 289)	(<i>n</i> = 76)	
Mean pre-enrollment unemployment rate	5.4%	5.4%	0.01
Mean pre-enrollment labor force	20,700	21,043	-0.05
Mean pre-enrollment average hourly earnings	\$29.61	\$29.56	0.09
Mean pre-enrollment average hours worked in a week	33.5	33.5	-0.11
Mean pre-enrollment minimum wage	\$9.48	\$9.45	0.05

 Table A4. Baseline Equivalence of Treatment and Control Groups, Benchmark Sample

⁶² See footnote 106.

BEAM IMPACT EVALUATION



Age	1	•••	•••		•••	••		۰.	ļ.	•	•••		. ļ	•	•••	•••	• •	•	••	••	•••	••	•••
Ethnicity	••••	•••	•••		••	••	• •	••	.	•	•			•	••	•••	• •	•	••	••	•••	••	••
White	••••	•••	•••		••	••	• •	• •		•	•••		-	•	••	•••	• •	•	••	••	••	••	•••
Black	••••	•••	•••	•••	••	••	• •	••	י רי	•	•••	•	• - - -	•	••	•••	• •	•	••	••	•••	•••	••
Multiple Races	••••	•••	•••	•••	••	••	• •	••	• •	•	••	•		٥.	••	•••	• •	•	••	••	••	••	••
Other	••••	•••	•••	•••	••	••	• •	• •	• •	•	•0•		•	•	••	•••	• •	•	••	••	••	••	••
Female		• • •	• • •	• • •	••	••	• •	• 0	·ŀ	•	••	•••	•	•	••	••	• •	•	••	••	••	••	••
Disability	••••	• • •	• • •	•••	••	••	• •	••	• •	•	••		•	•	••	•••	•	•	••	••	•••	••	••
HS or GED	••••	• • •	•••	•••	••	••	•••	••	۰k	, .	••		•	•	••	••	• •	•	••	••	••	••	••
Some College	••••	• • •	•••	•••	••	••	•••	••		•	••	•	,	•	••	•••	• •	•	••	••	••	••	••
College Degree		• • •	•••	•••	••	••	• •	• •	. ļ	•	••			•	••	•••	• •	•	••	••	•••	••	••
Low Income Status	••••	• • •	•••	•••	••	••	• •	••	·¦	••	•••	• •		•	••	•••	• •	•	•••	••	••	••	••
SSI or SSDI	••••	•••	•••	•••	••	••	• •	••	÷	0	•••		÷	•	••	•••	• •	•	••	••	•••	•••	••
TANF	••••	•••	•••	•••	••	••	• •	••	٠ģ	•	•••			•	••	•••	• •	•	••	•••	•••	•••	••
WIOA Adult	••••	• • •	• • •	•••	••	••	••	••	• •	•	••	•••		•	••	••	• •	•	••	••	••	••	••
WIOA Dislocated	••••	• • •	• • •	•••	••	••	••	••	·ŀ	•	••	۰ و	•	•	••	••	• •	•	••	••	••	••	••
WIOA Youth	••••	• • •	• • •	•••	••	• •	••	• •	·ŀ	•	••	•••	۰ŀ	•	• •	• 0	• •	•	••	••	••	••	••
Average Employment Status	••••	• • •	• • •	•••	••	••	•••	••	• •	•	••	•	••	•	••	••	• •	•	••	••	••	••	••
Employed Pre Q1	••••	• • •	• • •	•••	••	••	• •	••	. .	•	••	••	•	•	• •	•••	• •	•	••	••	••	••	••
Employed Pre Q2	••••	• • •	• • •	•••	••	••		••		•	••	• 0	, , ,	•	• •	•••	• •	•	••	••	••	••	••
Employed Pre Q3	••••	• • •	•••	•••	••	••	• •	••	.	•	•••	• •	•	•	••	•••	• •	•	••	••	••	••	••
Employed Pre Q4	••••	• • •	•••	•••	••	••	• •	••	•	•	•••	• •	⊳¦	•	••	•••	• •	•	••	••	•••	••	••
Average Wages	••••	• • •	• • •	•••	••	••	• •	••		•	•••	• 0		•	••	•••	• •	•	••	••	•••	••	••
Wages Pre Q1	••••	• • •	• • •	•••	••	••	• •	••	• •	•	•	•••		• •	••	••	• •	•	••	••	•••	•••	••
Wages Pre Q2	••••	• • •	•••	•••	••	••	••	••	·ŀ	•	• 0	•••	•	•	••	•••	• •	•	••	••	••	••	••
Wages Pre Q3	••••	• • •	• • •	•••	••	••	••	••	·ŀ	•	••	•••	•	•	• •	••	• •	•	••	••	••	••	••
Wages Pre Q4	••••	• • •	• • •	•••	••	••	• •	• •	• •	·	••	•••	•	•	• •	•••	• •	•	••	••	••	••	••
Average Hourly Earnings	••••	• • •	• • •	•••	••	••	• •	• •	• •	·	• •		•	•	• •	•••	• •	•	••	••	••	••	••
Average Weekly Hours	••••	•••	•••	•••	••	••		••	••	•	••	 	• I.	•	• •	•••	• •	•	••	••	•••	••	••
Labor Force	••••	• • •	•••	•••	••	••	• •	••	. .	•	•••			•	••	•••	• •	•	••	••	••	••	••
Minimum Wage	••••	• • •	• • •	•••	••	••	• •	••	·¦	•	0.	• •		•	••	•••	• •	•	•••	••	•••	•••	••
Unemployment Rate	1	• • •	• • •	•••	••	••	• •	••	÷	·	o ••		÷	•	•••	•••	• •	•	•••	•••	••	•••	••
				_		_			_									-					
				ا 6	। 5 ·	ا 4	। 3	ا 2	 1).	I 1 .	2	 .3	ا 4.		5	۱ .6					

Overall the benchmark treatment and control groups appear to be balanced on most characteristics at baseline. A well-executed randomization procedure does not guarantee balance on any given characteristic; what it does guarantee is that these features will be independent of the assignment to treatment or comparison condition. The Evaluation Team presents baseline equivalence statistics as a descriptive feature and a means to identify any irregularities. Standardized differences between the treatment and control groups are less than 0.25 for all but four characteristics – disability status, identifying as Multiracial, employed during the 3rd quarter prior to study entry, and enrolling as a youth participant under WIOA. The proportion of participants in the treatment group who have a disability is about 16% higher than the proportion of participants in the control group with a disability. Similarly, the proportion of participants in the control group members were about 10% more likely to be employed during the 3rd quarter prior to the quarter of study entry. The proportion of participants in the control group who identify as Multiracial is about 2% lower than the proportion in the treatment group.

An imbalance in key characteristics, such as disability status, between the two groups could suggest that assignment was nonrandom, either by design or a lack of fidelity. The Evaluation Team implemented a number of safeguards to prevent tampering by program staff and check the results of assignment throughout the duration of the study period. During these regular reviews, there was no evidence to suggest that assignment was nonrandom; however, the administrative data collection schedule from NYSDOL did not allow for ongoing monitoring of balance among key characteristics.

Table A5 below presents the treatment and control group means for each characteristic and the balance statistic in the form of standardized differences for the subsample of study participants enrolled before December 2017 and used in secondary (full-year subsample) analyses. Figure A2 presents the standardized mean differences graphically.

BEAM IMPACT EVALUATION

Characteristic	Treatment	Control	Standardized Difference
Age	(<i>n</i> = 269)	(<i>n</i> = 72)	
Mean age in years at enrollment	29.6	31.4	-0.21
Gender	(<i>n</i> = 248)	(<i>n</i> = 71)	
Female	62.9%	69.0%	-0.16
Race	(<i>n</i> = 235)	(<i>n</i> = 69)	
White	66.0%	73.9%	-0.24
Black	26.0%	20.3%	0.20
Multiracial	6.4%	4.3%	0.25
Other	1.7%	1.4%	0.07
Ethnicity	(<i>n</i> = 203)	(<i>n</i> = 59)	
Hispanic/Latino	10.3%	10.2%	0.02
Low income status	(<i>n</i> = 269)	(<i>n</i> = 72)	
Yes	53.9%	52.8%	0.01
Receives SSI or SSDI	(<i>n</i> = 269)	(<i>n</i> = 72)	
Yes	5.9%	5.6%	0.04
Receives TANF	(<i>n</i> = 269)	(<i>n</i> = 72)	
Yes	10.8%	12.5%	-0.14
Individual with a disability	(<i>n</i> = 239)	(<i>n</i> = 65)	
Yes	31.8%	15.4%	0.57
WIOA recipient	(<i>n</i> = 269)	(<i>n</i> = 72)	
Adult	76.6%	77.8%	-0.05
Youth	8.6%	4.2%	0.47
Dislocated worker	27.1%	23.6%	0.11
Highest educational level	(<i>n</i> = 269)	(<i>n</i> = 72)	
HS Diploma or equivalent	34.9%	38.9%	-0.12
Some college	52.4%	44.4%	0.20
Associate's Degree or higher	12.6%	16.7%	-0.19
Proportion employed pre-enrollment	(<i>n</i> = 269)	(<i>n</i> = 72)	
1 st quarter	58.7%	54.2%	0.11
2 nd quarter	61.7%	56.9%	0.12
3 rd quarter	60.6%	51.4%	0.23
4 th quarter	60.6%	52.8%	0.19
Mean quarterly wages pre-enrollment	(<i>n</i> = 269)	(<i>n</i> = 72)	
1 st quarter	\$2,535.58	\$2,380.10	0.05
2 nd quarter	\$2,705.90	\$2,599.93	0.03
3 rd quarter	\$2,904.55	\$2,255.14	0.18
4 th quarter	\$3,058.99	\$2,276.68	0.21
Regional labor statistics	(<i>n</i> = 269)	(<i>n</i> = 72)	
Mean pre-enrollment unemployment rate	5.4%	5.4%	0.01
Mean pre-enrollment labor force	20,733	21,272	-0.07
Mean pre-enrollment average hourly earnings	\$29.55	\$29.51	0.07
Mean pre-enrollment average hours worked in a week	33.5	33.5	-0.11
Mean pre-enrollment minimum wage	\$9.41	\$9.40	0.02

Table A5. Baseline Equivalence of Treatment and Control Groups, Full-Year Subsample

BEAM IMPACT EVALUATION



Age Image: Constraint of the second seco																														
Ethnicity White Black Multiple Races Other Female Disability HS or GED Some College College Degree Low Income Status SSI or SSDI TANF WIOA Adult WIOA Adult WIOA Adult WIOA Adult WIOA Adult WIOA Youth Average Employed Pre Q1 Employed Pre Q2 Employed Pre Q2 Employed Pre Q4 Average Mages Wages Pre Q4 Average Hourly Earnings Average Weekly Hours Labor Force Minimum Wage Unemployment Rate	Age	1		••	•••	• •		•••	• •	•	••		۰.		•			••	<u> </u>	•	•••	•	•••	•	••	•••	•••	•	•••	•
White Black Multiple Races Other Other Image: Construction of the second se	Ethnicity		•••	•••	•••	• •	•••	••	• •	•	••		••		•	o٠	ļ.	•••	 	•	•••	•	•••	•	••	•••	•••		•••	•
Black Multiple Races Other Female Disability HS or GED Some College College Degree Low Income Status SSI or SSDI TANF WIOA Adult WIOA Adul	White		•••	•••	•••	• •	•••	••	• •	•	••	••	• •		•			•••		•	•••	•	•••	•	••	••	•••	•	•••	•
Multiple Races Other Female Disability HS or GED Some College College Degree Low Income Status SSI or SSDI TANF WIOA Adult WIOA Adult WIOA Dislocated WIOA Youth Average Employment Status Employed Pre Q1 Employed Pre Q2 Employed Pre Q2 Employed Pre Q4 Average Wages Pre Q4 Wages Pre Q2 Wages Pre Q2 Wages Pre Q2 Wages Pre Q4 Average Wages Average Wages Average Wages Average Wages Average Wages Average Wages Average Wages Average Wages Average Wages Average Wages Multiple Races Minimum Wage Unemployment Rate WiOA Autor Minimum Wage Unemployment Rate Average Wages Multiple Races Minimum Wage Unemployment Rate Multiple Races Multiple Races	Black		• • •	••	•••	• •	••	••	• •	•	••	-	••	- -	• •		1	• • •	 • ·	• •	•••	•	••	•	••	•••	•••	•	••	•
Other Image: Constraint of the second se	Multiple Races		•••	••	••	• •	•••	••	• •	•	••	•	••	•	•	•	 	• •(•	••	•	••	•	••	••	•••	•	•••	•
Female Image: Previous of the second sec	Other		•••	••	• •	• •	••	••	• •	•	••	•	••	٠ŀ	•	••	١	• •	• ·	•	•••	•	••	•	••	••	• •	•	•••	•
Disability HS or GED Some College College Degree Low Income Status SSI or SSDI TANF WIOA Aduit WIOA Aduit WIOA Dislocated WIOA Youth Average Employed Pre Q1 Employed Pre Q2 Employed Pre Q2 Employed Pre Q4 Average Wages Wages Pre Q2 Wages Pre Q4 Average Weekly Hours Labor Force Minimum Wage Unemployment Rate	Female		•••	••	••	• •	••	••	• •	•	••	·	• •	• •	•	•	•	••	· ·	• •	••	•	••	•	••	••	••	•	••	•
HS or GED Some College College Degree Low Income Status SSI or SSDI TANF WIOA Adult WIOA Dislocated WIOA Youth Average Employment Status Employed Pre Q1 Employed Pre Q2 Employed Pre Q2 Employed Pre Q3 Employed Pre Q4 Average Wages Wages Pre Q1 Wages Pre Q2 Wages Pre Q1 Wages Pre Q2 Wages Pre Q2 Wages Pre Q3 Wages Pre Q4 Average Weekly Hours Labor Force Minimum Wage Unemployment Rate	Disability		•••	••	••	• •	••	••	• •	•	••	•	••	٠ŀ	•	•	•	• •	 .	•	••	•	• 0	•	••	••	• •	•	•••	•
Some College College Degree Low Income Status SSI or SSDI TANF WIOA Adult WIOA Dislocated WIOA Youth Average Employed Pre Q1 Employed Pre Q2 Employed Pre Q2 Employed Pre Q3 Employed Pre Q4 Average Wages Wages Pre Q1 Wages Pre Q2 Wages Pre Q2 Wages Pre Q2 Wages Pre Q3 Average Weekly Hours Labor Force Minimum Wage Unemployment Rate	HS or GED		• • •	••	••	• •	••	••	• •	•	••	•	••	0	•	•	•	••	.	• •	••	•	••	•	••	••	••	•	•••	•
College Degree Low Income Status SSI or SSDI TANF WIOA Adult WIOA Adult WIOA Youth Average Employment Status Employed Pre Q1 Employed Pre Q2 Employed Pre Q3 Employed Pre Q4 Average Wages Wages Pre Q1 Wages Pre Q2 Wages Pre Q2 Wages Pre Q3 Wages Pre Q4 Average Hourly Earnings Average Weekly Hours Labor Force Minimum Wage Unemployment Rate	Some College	• •	•••	••	•••	• •	••	••	• •	•	••	.	••	. I.	•	•		•0•	 	•	••	•	••	•	••	•••	•••	•	••	•
Low Income Status SSI or SSDI TANF WIOA Adult WIOA Adult WIOA Youth Average Employed Pre Q1 Employed Pre Q2 Employed Pre Q2 Employed Pre Q3 Employed Pre Q4 Average Wages Wages Pre Q1 Wages Pre Q2 Wages Pre Q2 Wages Pre Q3 Wages Pre Q4 Average Hourly Earnings Average Weekly Hours Labor Force Minimum Wage Unemployment Rate	College Degree	• •	•••	••	•••	• •	••	••	• •	•	•••	.	• •	. 	•	•	!.	••	<u>.</u> 	•	•••	•	••	•	••	•••	•••	•	••	•
SSI or SSDI TANF WIOA Adult WIOA Adult WIOA Youth Average Employment Status Employed Pre Q1 Employed Pre Q2 Employed Pre Q3 Employed Pre Q4 Average Wages Wages Pre Q1 Wages Pre Q2 Wages Pre Q2 Wages Pre Q3 Wages Pre Q4 Average Hourly Earnings Average Weekly Hours Labor Force Minimum Wage Unemployment Rate	Low Income Status		•••	•••	••	• •	••	••	•••	•	••		••		•	٥·	!. 	••	 	•	•••	•	••	•	••	••	••	•	•••	•
TANF WIOA Adult WIOA Adult WIOA Dislocated WIOA Youth Average Employment Status Employed Pre Q1 Employed Pre Q2 Employed Pre Q3 Employed Pre Q4 Average Wages Wages Pre Q1 Wages Pre Q2 Wages Pre Q3 Wages Pre Q4 Average Wages Average Wages Wages Pre Q4 Average Wages Unemployment Rate	SSI or SSDI		•••	•••	•••	• •	••	••	• •	•	•••	i	••	ij	•	•	÷	••	ŀ	•	•••	•	•••	•	••	•••	•••	•	••	•
WIOA Adult WIOA Dislocated WIOA Youth WIOA Youth Average Employment Status Image: Status Employed Pre Q1 Image: Status Employed Pre Q2 Image: Status Employed Pre Q3 Image: Status Employed Pre Q4 Image: Status Average Wages Image: Status Wages Pre Q1 Image: Status Wages Pre Q2 Image: Status Wages Pre Q3 Image: Status Wages Pre Q3 Image: Status Wages Pre Q3 Image: Status Wages Pre Q4 Image: Status Average Wages Image: Status Wages Pre Q3 Image: Status Wages Pre Q4 Image: Status Average Hourly Earnings Image: Status Average Weekly Hours Image: Status Labor Force Image: Status Minimum Wage Image: Status Image: I	TANF		•••	••	•••	• •	••	••	• •	•	•••	ľ	••	°	•	•	j.	••	ŀ	• •	•••	•	••	•	••	•••	•••	•	••	•
WIOA Dislocated WIOA Youth Average Employment Status Image: Constraint of the second status Employed Pre Q1 Image: Constraint of the second status Employed Pre Q1 Image: Constraint of the second status Employed Pre Q2 Image: Constraint of the second status Employed Pre Q3 Image: Constraint of the second status Employed Pre Q3 Image: Constraint of the second status Mages Pre Q3 Image: Constraint of the second status Wages Pre Q2 Image: Constraint of the second status Wages Pre Q3 Image: Constraint of the second status Wages Pre Q3 Image: Constraint of the second status Wages Pre Q3 Image: Constraint of the second status Wages Pre Q3 Image: Constraint of the second status Wages Pre Q3 Image: Constraint of the second status Wages Pre Q4 Image: Constraint of the second status Average Weekly Hours Image: Constraint of the second status Labor Force Image: Constraint of the second status Minimum Wage Image: Constraint of the second status Unemployment Rate Image: Constraint of the second status Image: Constraint of the second status Image: Consecond status <td>WIOA Adult</td> <td> • •</td> <td>•••</td> <td>••</td> <td>••</td> <td>• •</td> <td>•••</td> <td>••</td> <td>•••</td> <td>•</td> <td>••</td> <td>• </td> <td>••</td> <td>•</td> <td>0</td> <td>•</td> <td>•</td> <td>••</td> <td>•</td> <td>• •</td> <td>••</td> <td>•</td> <td>••</td> <td>•</td> <td>••</td> <td>••</td> <td>••</td> <td>•</td> <td>•••</td> <td>•</td>	WIOA Adult	• •	•••	••	••	• •	•••	••	•••	•	••	•	••	•	0	•	•	••	•	• •	••	•	••	•	••	••	••	•	•••	•
WIOA Youth WIOA Youth Average Employment Status Imployed Pre Q1 Employed Pre Q2 Imployed Pre Q2 Employed Pre Q3 Imployed Pre Q3 Employed Pre Q4 Imployed Pre Q4 Average Wages Imployed Pre Q2 Wages Pre Q1 Imployed Pre Q2 Wages Pre Q1 Imployed Pre Q3 Wages Pre Q1 Imployed Pre Q4 Average Wages Imployed Pre Q4 Average Pre Q3 Imployed Pre Q4 Wages Pre Q3 Imployed Pre Q4 Average Hourly Earnings Imployed Pre Q4 Average Weekly Hours Imployed Pre Q4 Labor Force Imployment Rate Unemployment Rate Imployed Pre Q4 Average Weekly Hours Imployed Pre Q4 Imployed Pre Q4 Imployed Pre Q4 Imployed Pre Q4 Imployed Pre Q4 Average Weekly Hours Imployed Pr	WIOA Dislocated	• •	•••	••	••	• •	•••	••	• •	•	••	•	••	·ŀ	• •	•	÷	••	•	• •	••	•	••	•	••	••	••	•	•••	•
Average Employment Status Employed Pre Q1 Employed Pre Q2 Employed Pre Q2 Employed Pre Q3 Employed Pre Q4 Average Wages Wages Pre Q1 Wages Pre Q2 Wages Pre Q2 Wages Pre Q3 Wages Pre Q4 Average Hourly Earnings Average Weekly Hours Labor Force Minimum Wage Unemployment Rate	WIOA Youth	• •	•••	••	• •	• •	••	••	•••	•	••		••	٠ŀ	•	•	۰ŀ	• •	• ·	• •	• •	0.	••	•	••	••	• •	•	••	•
Employed Pre Q1 Employed Pre Q2 Employed Pre Q3 Employed Pre Q4 Average Wages Wages Pre Q1 Wages Pre Q2 Wages Pre Q2 Wages Pre Q3 Wages Pre Q3 Wages Pre Q4 Average Hourly Earnings Average Weekly Hours Labor Force Minimum Wage Unemployment Rate	Average Employment Status	• •	•••	••	••	• •	••	••	• •	•	••	•	••	·	•	•	•	•••	. 	• •	••	•	••	٠	••	••	••	•	••	•
Employed Pre Q2 Employed Pre Q3 Employed Pre Q4 Average Wages Wages Pre Q1 Wages Pre Q2 Wages Pre Q3 Wages Pre Q3 Wages Pre Q4 Average Hourly Earnings Average Weekly Hours Labor Force Minimum Wage Unemployment Rate	Employed Pre Q1	• •	•••	••	••	• •	••	••	• •	•	••		••	• •	•	•	.	••	. .	•	••	•	••	•	••	••	••	•	••	•
Employed Pre Q3 Employed Pre Q4 Average Wages Wages Pre Q2 Wages Pre Q2 Wages Pre Q3 Wages Pre Q4 Average Hourly Earnings Average Weekly Hours Labor Force Minimum Wage Unemployment Rate	Employed Pre Q2	• •	•••	••	••	• •	••	••	•••	•	••	.	••	. l. 	•	•	. ₀ 	••	I 	• •	••	•	••	•	••	••	••	•	•••	•
Employed Pre Q4 Average Wages Wages Pre Q1 Wages Pre Q2 Wages Pre Q3 Wages Pre Q4 Average Hourly Earnings Average Weekly Hours Labor Force Minimum Wage Unemployment Rate	Employed Pre Q3		•••	•••	•••	• •	•••	•••	•••	•	••		••		•	•	.! . 	• 0	. 	• •	•••	•	•••	•	••	••	•••	•	•••	•
Average Wages Wages Pre Q1 Wages Pre Q2 Wages Pre Q2 Wages Pre Q3 Wages Pre Q3 Wages Pre Q4 Wages Pre Q4 Average Hourly Earnings Image Pre Q4 Average Weekly Hours Image Pre Q4 Labor Force Image Pre Q4 Minimum Wage Image Pre Q4 Unemployment Rate Image Pre Q4 Average Weekly Hours Image Pre Q4 Average Weekly Hours Image Pre Q4 Labor Force Image Pre Q4 Minimum Wage Image Pre Q4 Image Pre Q4 Image Pre Q4 Average Weekly Hours Image Pre Q4 Image Pre Q4	Employed Pre Q4		•••	•••	•••	• •	•••	••	•••	•	•••	ť	••	÷	•	••	÷	••	¦. 	•	•••	•	•••	•	••	••	•••	•	•••	•
Wages Pre Q1 Wages Pre Q2 Wages Pre Q3 Wages Pre Q4 Average Hourly Earnings Average Weekly Hours Labor Force Minimum Wage Unemployment Rate	Average Wages	•••	•••	••	••	• •	••	••	•••	•	•••	Ì	••	Ì	•	•	•	• •	i I	•	••	•	••	•	••	••	••	•	••	•
Wages Pre Q2 Wages Pre Q3 Wages Pre Q4 Average Hourly Earnings Average Weekly Hours Labor Force Minimum Wage Unemployment Rate 654321 0 .1 .2 .3 .4 .5 .6	Wages Pre Q1		•••	•••	•••	• •	•••	•••	• •	•	•••	•	••	•	• •	•	•	•••	•	• •	•••	٠	•••	•	••	•••	•••	•	•••	•
Wages Pre Q3 Wages Pre Q4 Average Hourly Earnings Average Weekly Hours Labor Force Minimum Wage Unemployment Rate 654321 0 .1 .2 .3 .4 .5 .6	Wages Pre Q2	• •	•••	••	•••	• •	•••	••	•••	•	••	•	••	·ŀ	• •	0	•	••	 • ·	• •	••	•	••	•	••	••	•••	•	•••	•
Wages Pre Q4 Average Hourly Earnings Average Weekly Hours Labor Force Minimum Wage Unemployment Rate 654321 0 .1 .2 .3 .4 .5 .6	Wages Pre Q3	• •	•••	••	••	• •	••	••	• •	•	••	•	••	•	•	•	•	• 0	• ·	• •	••	•	••	•	••	••	••	•	•••	•
Average Hourly Earnings Average Weekly Hours Labor Force Minimum Wage Unemployment Rate 6	Wages Pre Q4		• • •	••	•••	• •	• •	••	• •	•	••	•	••	٠ŀ	•	•	•	• •	• ·	•	••	٠	••	•	••	••	•••	•	••	•
Average Weekly Hours	Average Hourly Earnings		• • •	••	•••	• •	••	••	• •	•	••	•	••	• •	•	••		••	 	•	••	•	••	•	••	••	•••	•	••	•
Labor Force Minimum Wage Unemployment Rate 654321 0 .1 .2 .3 .4 .5 .6	Average Weekly Hours		•••	••	••	• •	••	••	•••	•	••	.	••	•0•	•	•		••	 	• •	••	٠	••	•	••	••	••	•	•••	•
Minimum Wage Unemployment Rate 654321 0 .1 .2 .3 .4 .5 .6	Labor Force		•••	•••	•••	• •	•••	••	•••	•	••	.	••		•••	•	. . 	••	۱ <u>.</u> ا	• •	•••	•	••	•	••	••	•••	•	•••	•
Unemployment Rate	Minimum Wage	··	•••	•••	•••	• •	••	••	• •	•	•••	•	••	-	•	٥.	¦. 	••	". 	•	•••	•	•••	•	••	•••	•••	•	•••	•
654321 0 .1 .2 .3 .4 .5 .6	Unemployment Rate	1	• • •	••	••	• •	••	••	• •	•	•••	i	••	ij	•	b٠	į.	••	i.	• •	••	•	••	•	••	••	••	•	•••	•
654321 0 .1 .2 .3 .4 .5 .6										_			_	_			1		L			_		1						
							/	6 -	.5	ا 4	ا ا	3	2		1	" 0	.1	۱ 2.	1	3.	۱ .4	۱ 5.	; .	۱ .6						

Similar to the benchmark sample, the secondary subsample treatment and control groups appear to be balanced on most characteristics at baseline overall. Standardized mean differences between the treatment and control groups are less than 0.25 for all but two characteristics – disability status and enrolling as a youth participant under WIOA.

BENCHMARK ANALYTIC RESULTS

MODEL SELECTION

As explained in the Analytic Model Specification section, the preferred benchmark model was prespecified by the Evaluation Team in the EDR prior to receiving any data. To determine whether the benchmark approach was the best fitting model, the Evaluation Team conducted a series of goodness of fit and likelihood ratio tests using the below specified models:

- Model 1: Regress outcome on treatment indicator (equivalent to a t-test).
- Model 2: Regress outcome on treatment indicator and blocking variables.
- **Model 3**: Regress outcome on treatment indicator, blocking variables, and pre-intervention measure of the outcome.
- **Model 4** (Benchmark): Regress outcome on treatment indicator, blocking variables, preintervention outcome measure, and a limited set of covariates.⁶³
- **Model 5**: Regress outcome on treatment indicator, blocking variables, pre-intervention outcome measure, and a comprehensive set of covariates.⁶⁴

The results of goodness of fit tests are presented in Table A6 below. Although Model 1 appears to be the best fitting model for outcomes within the education domain and Model 3 appears to be the best fit for employment domain outcomes, the Evaluation Team retained Model 4 as the benchmark approach because it was pre-specified in the EDR. The impact estimates between the five models are essentially identical.

⁶³ The benchmark model includes the following set of covariates: highest level of education, age at enrollment, gender, race, ethnicity, low income status, and disability status.

⁶⁴ Model 5 includes the covariates used in the benchmark model, with the addition of the following: homeless status, SSI/SSDI recipient, TANF recipient, UI recipient, WIOA funding (adult, dislocated worker, youth).

Table A6. Goodness of Fit Test Results (n = 365) Content

	Model 1	Model 2	Model 3	Model 4	Model 5
				(Benchmark)	
Research Question 1					
AIC	384.96	389.38	389.70	396.94	402.37
BIC	392.76	428.37	436.50	490.54	523.27
Likelihood-ratio test (df)		11.58(8)	3.67(2)	16.76(12)	8.57(7)
Research Question 2					
AIC	1113.17	1113.97	1116.42	1126.14	1129.38
BIC	1120.97	1152.97	1163.21	1219.73	1250.28
Likelihood-ratio test		15.20(8)~	1.56(2)	14.28(12)	10.75(7)
Research Question 3					
AIC	-249.62	-245.99	-244.28	-238.17	-229.64
BIC	-241.82	-206.99	-197.48	-144.57	-108.75
Likelihood-ratio test		12.36(8)	2.29(2)	17.89(12)	5.48(7)
Research Question 4					
AIC	474.83	466.67	405.65	419.73	427.42
BIC	482.63	505.67	448.55	509.43	544.41
Likelihood-ratio test		24.16(8)**	63.02(1)***	9.92(12)	6.32(7)
Research Question 5					
AIC	7006.48	7004.55	6944.70	6959.95	6967.27
BIC	7014.28	7043.55	6987.60	7049.65	7084.26
Likelihood-ratio test		17.93(8)*	61.85(1)***	8.75(12)	6.69(7)
Research Question 6					
AIC	1228.62	1226.39	1145.01	1153.99	1157.94
BIC	1236.42	1265.38	1187.91	1243.69	1274.94
Likelihood-ratio test		18.23(8)*	83.37(1)***	15.02(12)	10.05(7)

Note: *** *p* < 0.001, ** *p* < 0.01, * *p* < 0.05, ~ *p* < 0.10.

IMPACT ESTIMATES

BENCHMARK IMPACT ESTIMATES

The results of the benchmark analytic model for each of the six research questions are presented in Table A7 below.

 Table A7. Benchmark Analytic Model Results (n = 365)

Estimate	Standard Error	Effect Size
0.03	0.05	0.11
0.11	0.15	0.10
0.01	0.02	0.15
0.00	0.06	0.00
\$126.23	\$434.20	0.04
-0.03	0.15	-0.03
	Estimate 0.03 0.11 0.01 0.00 \$126.23 -0.03	Estimate Standard Error 0.03 0.05 0.11 0.15 0.01 0.02 0.00 0.06 \$126.23 \$434.20 -0.03 0.15

Note: *** p < 0.001, ** p < 0.01, * p < 0.05, ~ p < 0.10.

FULL-YEAR SUBSAMPLE IMPACT ESTIMATES

In addition to the primary study for each research question, the Evaluation Team conducted a series of secondary analyses that limited the analytic sample to individuals enrolled in the Impact Study by December 31, 2017. This was the originally planned cutoff for enrollment (documented in the EDR) so that individuals would have a full year of follow-up time in order to achieve outcomes.

 Table A8. Full-Year Subsample Analytic Model Results (n = 341)

	Estimate	Standard Error	Effect Size
Educational Outcomes			
Enrollment in Postsecondary Program at 1 Year			
Estimate of program impact	0.08	0.06	0.27
Number of Consecutive Semesters			
Estimate of program impact	0.14	0.15	0.13
Graduated			
Estimate of program impact	0.01	0.02	0.23
Employment Outcomes			
Employed During 4 Quarters Post			
Estimate of program impact	-0.03	0.06	-0.08
Average Wages in 4 th Quarter Post			
Estimate of program impact	-\$131.65	\$447.34	-0.04
Number of Quarters Employed Out of 4 Quarters			
Estimate of program impact	-0.04	0.21	-0.02

Note: *** *p* < 0.001, ** *p* < 0.01, * *p* < 0.05, ~ *p* < 0.10.

Additional discussion of the results of the secondary (full-year subsample) analyses on employment outcomes is provided below.

RESEARCH QUESTION 4: IMPACT ON EMPLOYMENT FOR SUBSAMPLE

Table A8 presents the impact estimate on the likelihood that a subsample participant would be employed during any of the four full quarters after study entry. Similar to the benchmark findings, the secondary study indicates that the GCP program did not have a statistically significant impact on participants' likelihood of becoming employed during any post-enrollment quarters.

Figure A3 presents the unadjusted proportion of individuals in the secondary subsample who were employed during four pre-enrollment quarters, quarter of study entry, and four post-enrollment quarters, by treatment condition.



Figure A3. Percent of Full-Year Subsample Participants Employed During Pre- and Post-Enrollment Quarters (n = 341)⁶⁵

Benchmark results indicate that the GCP intervention had no impact on participants' likelihood of becoming employed during any of the three quarters after study entry. Findings produced by the secondary analysis are substantively identical. Both lines demonstrate that post-program employment for both conditions is slightly better than it was at the time of enrollment, but not dramatically so. The lines also demonstrate the similarity between the two groups' employment trends before and after study enrollment. The uncertainty of difference in the outcome is illustrated by the crisscrossing of the two lines post enrollment; At each time point there are small differences between the two groups; however those differences are in flux (e.g., a positive difference at post quarter 1 becomes a negative difference at post quarter 2). This inconstancy could be motivated by actual differences, but with the small comparison sample, it is likely imprecision driven by measurement error and the natural variation in human experience.

RESEARCH QUESTION 5: IMPACT ON QUARTERLY WAGES

Table A8 presents the impact estimate on the total wages earned during the fourth full quarter after study entry for this subsample. Consistent with the benchmark findings, the secondary study indicates that the GCP program did not have a statistically significant impact on participants' total wages during the 4th quarter after entry.

Figure A4 presents the unadjusted wages earned during four pre-enrollment quarters, quarter of study entry, and four post-enrollment quarters, by treatment condition, for the subsample included in the secondary analysis. CCS participants generally earned more wages during the four quarters after study entry than their GCP counterparts.

⁶⁵ Figure A3 presents the unadjusted means for the percentage of participants who were employed during each of the four quarters prior to study enrollment, the quarter of study enrollment, and up to four quarters after study enrollment. Analytic sample size is 341 for all time points.



*Figure A4. Full-Year Subsample Average Wages Earned During Pre- and Post-Enrollment Quarters (n = 341)*⁶⁶

Consistent with the benchmark study, though both groups of participants improved their average wages after study entry, the secondary analysis indicates that even with more time, the impact of the program did not increase.

RESEARCH QUESTION 6: IMPACT ON EMPLOYMENT PERSISTENCE

Results from the secondary analysis of Research Question 6 (employment persistence) are presented in Table A8. For this subgroup, GCP participants were employed for, on average, 2.3 quarters, whereas CCS participants were employed for 2.4 quarters.

Similar to the benchmark findings, the secondary analysis showed that, even with more time, the impact of the GCP program on employment persistence remained the same. The standardized effect size for the impact on persistence is again close to zero.

UNEMPLOYED SUBGROUP IMPACT ESTIMATES

In addition to the benchmark and secondary analyses described in the Impact Study Findings section, the Evaluation Team conducted a third, exploratory subgroup analysis that limited the analytic sample to individuals who were unemployed during the quarter that directly preceded study entry, to examine whether there were any meaningful variations in program impact on outcomes compared with the full sample. Table A9 presents the results of these additional analyses, which are briefly discussed in the Discussion section.

⁶⁶ Figure A4 presents the unadjusted means for the quarterly earned wages during each of the four quarters prior to study enrollment, the quarter of study enrollment, and four quarters after study enrollment. Analytic sample size is 341 for all time points.

BEAM IMPACT EVALUATION

Table A9. Unemployed Subgroup Analytic Model Results (n = 149)

	Estimate	Standard Error	Effect Size
Educational Outcomes			
Enrollment in Postsecondary Program at 1 Year			
Estimate of program impact	0.03	0.08	0.11
Number of Consecutive Semesters			
Estimate of program impact	-0.04	0.20	-0.04
Graduated			
Estimate of program impact	-0.01	0.04	-0.09
Employment Outcomes			
Employed During 4 Quarters Post			
Estimate of program impact	0.11	0.10	0.27
Average Wages in 4 th Quarter Post			
Estimate of program impact	\$428.42	\$639.52	0.14
Number of Quarters Employed out of 4 Quarters			
Estimate of program impact	0.35	0.25	0.29

Note: *** p < 0.001, ** p < 0.01, * p < 0.05, ~ p < 0.10.

Impact estimates in Table A9 show that the GCP program did not have any discernible impact on educational outcomes for the subgroup of participants who were unemployed during the quarter that directly preceded the quarter of study enrollment. In terms of employment outcomes, however, the results of this exploratory study provide some evidence to suggest that the GCP program may benefit this unemployed subgroup compared with the CCS program.

The GCP program did not have any discernible effect on employment outcomes for the full (benchmark) and secondary (full-year subsample) analytic samples. Impact estimates and standardized effect sizes are close to zero in both cases (as shown in Tables B7 and B8). Estimates presented in Table A9 show that when the sample is restricted to the subgroup of participants who were unemployed directly prior to study entry (n = 149), impact estimates increase for all three employment outcomes (though remain not significant) and standardized effect sizes increase to 0.3 for Research Questions 4 (employed) and 6 (employment persistence). Although the results of these exploratory analyses cannot be considered confirmatory, they do provide some additional context for GCP's impact on participants' employment outcomes, namely that the program may benefit individuals who are unemployed at entry more so compared with the general study population.

APPENDIX B: IMPACT EVALUATION SENSITIVITY STUDIES

DESCRIPTION OF SENSITIVITY STUDIES

In order to test the extent to which the benchmark findings are robust to alternative assumptions and analytic specifications, the Evaluation Team conducted several additional analyses for each outcome. The benchmark findings reflect the preferred design and analytic approach, the one the Evaluation Team believes to be both the most defensible and simple to interpret. These additional sensitivity studies included alternative modeling and measurement approaches that were discarded because they did not fit the data as well, incorporated unmet assumptions, were too complex, or would not work with the full analytic sample. The results of these studies are not reported in detail; this would unnecessarily complicate the presentation of the results. Nevertheless, these studies have been conducted and are included in the consideration of the findings as a means of testing the extent to which the benchmark results persist. The benchmark approach is outlined in the Analytic Model Specifications section. Below is a brief overview of the sensitivity studies that were performed to test this approach:

- 1. Run the benchmark model with covariates that did not contain imputed values.
- 2. Run a reduced form of the benchmark model with only blocking covariates and without demographic covariates.
- 3. Run the benchmark model with the addition of regional time-varying economic and contextual variables (e.g., local unemployment rate, average weekly hours worked, etc.).
- 4. For Research Questions 2 and 6, run a series of regression models designed for count data (e.g., Poisson, negative binomial, zero-inflated negative binomial) on the outcome.
- 5. For Research Questions 1, 3, and 4, run a logit model on the dichotomous outcome.
- 6. For Research Questions 1, 3, and 4, run a multilevel logit model with random effects for blocking covariates.
- 7. For Research Question 3, run a Firth logit model as the probability of the outcome is extremely small (e.g., <10%).
- 8. For Research Question 1, construct an outcome that measures whether a participant ever enrolled in postsecondary training by the end of the study period. Follow-up periods vary for each participant.
- 9. For Research Question 5, run the benchmark model with inflation-adjusted wages.
- 10. For Research Questions 4 and 6, include the quarter in which the participant enrolled in the study (e.g., post-quarter 0) in the calculation of the outcome.

Table B1 provides a visual representation of the types of sensitivity analyses that were conducted for each research question.

Alternative Model	RQ1	RQ2	RQ3	RQ4	RQ5	RQ6
1. Covariates without imputed values	\checkmark	~	~	✓	~	~
2. Blocking covariates only	\checkmark	\checkmark	\checkmark	\checkmark	\checkmark	✓
3. Benchmark model plus regional economic variables	✓	✓	✓	✓	✓	~
4. Count models		~				✓
5. Logit model	\checkmark		✓	✓		
6. Multilevel logit model	✓		\checkmark	✓		
7. Firth logit model			✓			
8. Outcome defined as enrolled by end of study	\checkmark					
9. Inflation-adjusted wages					✓	
10. Include quarter of enrollment in calculation				✓		~

Table B1. Sensitivity Analyses Conducted, by Research Question

SENSITIVITY STUDY ANALYTIC RESULTS

As described above, the Evaluation Team conducted a series of sensitivity analyses for each of the six outcomes in the Impact Study to test the robustness of the benchmark approach to alternative assumptions and analytic specifications. The results of these tests are presented for each research question below. Figures B1 through B6 present the treatment coefficients and their 95% confidence intervals for the benchmark analytic model (solid red line), the secondary subsample model (red dashed lined), and each additional sensitivity analysis conducted for the outcome (solid grey lines). Tables C2 through B7 present the impact estimates, standard error, sample size, test statistic, and p-value of each analysis.⁶⁷

As shown in the figures and tables below, the results of the sensitivity analyses are overall consistent with the benchmark approach. For all outcomes, the results are null; this is indicated by the horizontal lines that represent the 95% confidence intervals. In all cases, the value of zero is within the 95% range, indicating a test was not significant.

For the education outcomes, impact estimates are in the direction that was hypothesized (positive), but range in magnitude. For the employment outcomes, impact estimates are generally close to zero, and vary in their directionality.

⁶⁷ Results from nonlinear (e.g., logit, Poisson) models are not represented in Figures C1 through C6 but are represented in the accompanying tables.

Figure B1. Treatment Coefficients and 95% Confidence Intervals for Benchmark Analytic Model and Sensitivity Analyses of Research Question 1



Table B2. Results of Benchmark Analytic Model and Sensitivity Analyses of Research Question 1

	Estimate	Standard Error	N	Test Statistic(p)
Benchmark Model	0.03	0.05	365	0.52(0.60)
Full-year subsample	0.08	0.06	341	1.31(0.19)
Ever enrolled in postsecondary training by end of study	0.04	0.06	365	0.74(0.46)
Benchmark model covariates without imputed values	0.01	0.06	242	0.14(0.89)
Blocking covariates only	0.03	0.05	365	0.60(0.55)
Benchmark model plus regional economic variables	0.02	0.05	365	0.42(0.67)
Logit model	0.23	0.35	365	0.65(0.51)
Multilevel logit model	0.23	0.35	365	0.65(0.52)

Figure B2. Treatment Coefficients and 95% Confidence Intervals for Benchmark Analytic Model and Sensitivity Analyses of Research Question 2



 Table B3. Results of Benchmark Analytic Model and Sensitivity Analyses of Research Question 2

	Estimate	Standard Error	N	Test Statistic(p)
Benchmark Model	0.11	0.15	365	0.74(0.46)
Full-year subsample	0.14	0.15	341	0.92(0.36)
Benchmark model covariates without imputed values	0.10	0.17	242	0.57(0.57)
Blocking covariates only	0.10	0.14	365	0.69(0.49)
Benchmark model plus regional economic variables	0.12	0.15	365	0.86(0.39)
Poisson model	1.22	0.31	365	0.78(0.44)
Negative binomial model	0.16	0.26	365	0.59(0.56)
Zero-inflated negative binomial model	0.19	0.25	365	0.76(0.45)

Figure B3. Treatment Coefficients and 95% Confidence Intervals for Benchmark Analytic Model and Sensitivity Analyses of Research Question 3



Table B4. Results of Benchmark Analytic Model and Sensitivity Analyses of Research Question 3

	Estimate	Standard Error	N	Test Statistic(p)
Benchmark Model	0.01	0.02	365	0.31(0.76)
Full-year subsample	0.01	0.02	341	0.43(0.66)
Benchmark model covariates without imputed values	0.02	0.02	242	0.70(0.49)
Blocking covariates only	0.01	0.02	365	0.26(0.80)
Benchmark model plus regional economic variables	0.01	0.02	365	0.30(0.77)
Logit model	0.26	0.89	168	0.29(0.77)
Multilevel logit model	0.14	0.86	316	0.16(0.87)
Firth logit model	0.07	0.74	365	0.10(0.92)

Figure B4. Treatment Coefficients and 95% Confidence Intervals for Benchmark Analytic Model and Sensitivity Analyses of Research Question 4



Table B5. Results of Benchmark Analytic Model and Sensitivity Analyses of Research Question 4

	Estimate	Standard Error	N	Test Statistic(p)
Benchmark Model	0.00	0.06	365	0.00(1.00)
Full-year subsample	-0.03	0.06	341	-0.46(0.65)
Benchmark model covariates without imputed values	-0.04	0.07	242	-0.65(0.52)
Blocking covariates only	0.00	0.06	365	0.03(0.98)
Benchmark model plus regional economic variables	0.01	0.06	365	0.18(0.86)
Include quarter of enrollment in calculation of outcome (3 quarters post)	0.00	0.05	365	0.03(0.97)
Full-year subsample, include quarter of enrollment in calculation of outcome (4 quarters post)	-0.03	0.05	341	-0.57(0.57)
Logit model	0.02	0.33	361	0.07(0.95)
Multilevel logit model	-0.03	0.33	365	-0.10(0.92)

Figure B5. Treatment Coefficients and 95% Confidence Intervals for Benchmark Analytic Model and Sensitivity Analyses of Research Question 5



Table B6. Results of Benchmark Analytic Model and Sensitivity Analyses of Research Question 5

	Estimate	Standard Error	N	Test Statistic(p)
Benchmark Model	\$126.23	\$434.20	365	0.29(0.77)
Full-year subsample	-\$131.65	\$447.34	341	-0.29(0.77)
Benchmark model covariates without imputed values	\$92.90	\$553.38	242	0.17(0.87)
Blocking covariates only	\$145.28	\$453.32	365	0.32(0.75)
Benchmark model plus regional economic variables	\$176.69	\$435.41	365	0.41(0.69)
Wages adjusted for inflation	\$114.81	\$427.95	365	0.27(0.79)

Figure B6. Treatment Coefficients and 95% Confidence Intervals for Benchmark Analytic Model and Sensitivity Analyses of Research Question 6



Table B7. Results of Benchmark Analytic Model and Sensitivity Analyses of Research Question 6

	Estimate	Standard Error	N	Test Statistic(p)
Benchmark Model	-0.03	0.15	365	-0.23(0.82)
Full-year subsample	-0.04	0.21	341	-0.19(0.85)
Benchmark model covariates without imputed values	-0.11	0.18	242	-0.61(0.54)
Blocking covariates only	-0.02	0.17	365	-0.11(0.91)
Benchmark model plus regional economic variables	0.00	0.15	365	-0.01(0.99)
Include quarter of enrollment in calculation of outcome (up to 3 quarters post)	-0.11	0.19	365	-0.57(0.57)
Full-year subsample, include quarter of enrollment in calculation of outcome (up to 4 quarters post)	-0.11	0.24	341	-0.45(0.66)
Poisson count model	0.97	0.08	365	-0.31(0.75)
Negative binomial count model	-0.03	0.08	365	-0.31(0.75)
Zero-inflated negative binomial count model	-0.02	0.07	365	-0.32(0.75)