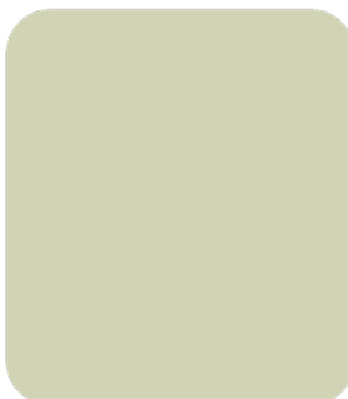


The San Diego Workforce Partnership's Bridge to Employment in the Healthcare Industry Program

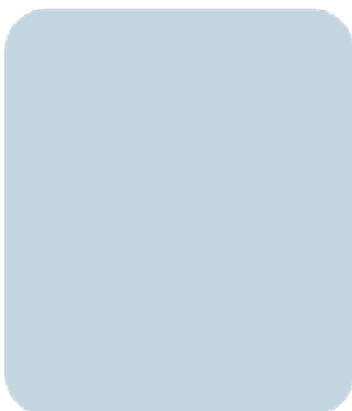


Appendices for Three-Year Impact Report



OPRE Report 2020-105

August 2020



PACE
Pathways for Advancing
Careers and Education

The San Diego Workforce Partnership's Bridge to Employment in the Healthcare Industry Program: Appendices for Three-Year Impact Report

A Pathways for Advancing Careers and Education (PACE) / Career Pathways Intermediate Outcomes Study Publication

OPRE Report 2020-105

August 2020

David Judkins, Samuel Dastrup, and Randall Juras, Abt Associates
Mary Farrell, MEF Associates

Submitted to:

Nicole Constance and Amelia Popham, Project Officers
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

Contract Number: HHSP23320095624WC, Task Order HHSP23337019T

Project Director: Larry Buron
Principal Investigator: David Fein; Director of Analysis: David Judkins
Abt Associates
6130 Executive Boulevard
Rockville, MD 20852

This report is in the public domain. Permission to reproduce is not necessary. Suggested citation: Judkins, David, Samuel Dastrup, Randall Juras, and Mary Farrell. 2020. *The San Diego Workforce Partnership's Bridge to Employment in the Healthcare Industry Program: Appendices for Three-Year Impact Report*. OPRE Report 2020-105. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Disclaimer: The views expressed in this publication do not necessarily reflect the views or policies of the Office of Planning, Research, and Evaluation, the Administration for Children and Families, or the U.S. Department of Health and Human Services.

This report and other reports sponsored by the Office of Planning, Research, and Evaluation are available at www.acf.hhs.gov/opre.



[Sign-up for the OPRE Newsletter](#)



[Follow OPRE
on Twitter
@OPRE_ACF](#)



[Like OPRE's page
on Facebook
OPRE.ACF](#)



[Follow OPRE
on Instagram
@opre_acf](#)



[Connect on
LinkedIn
company/opreacf](#)



Contents

Appendix A: Baseline Characteristics and Adjustments	1
A.1 Details on Baseline Covariates	1
A.2 Comparing Treatment and Control Groups at Baseline	4
A.3 Regression Adjustment.....	6
Appendix B: Three-Year Survey Data.....	15
B.1 Measures Based on Follow-up Survey Data	16
B.2 Imputation in the Three-Year Survey	22
B.3 Survey Nonresponse Analysis	35
B.4 Quality and Completeness of Exam-Based Credentials Reported in the Survey ...	46
B.5 Quality and Completeness of School-Issued Credentials Reported in the Survey	47
Appendix C: NDNH's Unemployment Insurance Wage Data	51
C.1 Data Collection Process.....	51
C.2 Data and Measures	52
Appendix D: Comparing NDNH- and Survey-Based Employment and Earnings Estimates	54
Appendix E: Treatment of Outliers	56
Appendix F: Cost-Benefit Analysis Supplement	57
F.1 Cost of Bridge to Employment Program and Control Group Alternatives.....	61
F.2 Cost of Education or Training.....	66
F.3 Earnings Impacts, Fringe Benefits, Taxes, and Means-tested Assistance.....	74
F.4 Uncertainty in Components of the Cost-Benefit Analysis	78
F.5 Data Sources	81
Appendix References	84

List of Exhibits

Exhibit A-1:	Operationalization of Baseline Measures Used as Covariates in Regression-Adjusted Impact Estimates.....	2
Exhibit A-2:	Baseline Balance	5
Exhibit A-3:	Covariates Selected, by Outcome Domain.....	12
Exhibit A-4:	Comparison of Confirmatory and Secondary Impact Estimates Unadjusted and Adjusted for Baseline Imbalances	14
Exhibit B-1:	Details on Specifications for Survey-Based Education Outcomes in Chapter 3.....	17
Exhibit B-2:	Details on Specifications for Survey-Based Employment/Earnings Outcomes in Chapter 4	18
Exhibit B-3:	Details on Specifications for Survey-Based Intermediate Outcomes in Chapter 4.....	19
Exhibit B-4:	Details on Specifications for Survey-Based Other Life Outcomes in Chapter 5	20
Exhibit B-5:	Details on Specifications for Survey-Based Child Outcomes in Chapter 5	21
Exhibit B-6:	Imputation Rates among Survey Respondents in the Evaluation Sample for Bridge to Employment.....	23
Exhibit B-7:	Comparison of Selected Impact Estimates of Bridge to Employment with and without Skipouts.....	31
Exhibit B-8:	Date Imputation for Three-Year Impact Study (Pooled Sample).....	34
Exhibit B-9:	Comparison of Selected Impact Estimates of Bridge to Employment with and without Imputation of NSC-Inferred Unreported Spells	35
Exhibit B-10:	Baseline Balance on Full Sample, Unweighted Respondent Sample, and Weighted Respondent Sample.....	38
Exhibit B-11:	Comparison of Selected Estimates of the Impact of Bridge to Employment for the Unweighted and Weighted Survey Samples	41
Exhibit B-12:	Comparison of Selected Impact Estimates on Credential Receipt in the First 18 Follow-up Months for Bridge to Employment	50
Exhibit D-1:	Impacts of Bridge to Employment on Earnings and Employment around Follow-up Q12 Based on Wage Records and Self-Reports.....	55
Exhibit F-1:	Costs and Benefits of Bridge to Employment, by Perspective	58
Exhibit F-2:	Sources Used to Estimate Per-Participant Cost of the Bridge to Employment Program	62
Exhibit F-3:	Reported Differences in Level of Services Accessed by Control Group Members.....	63

Exhibit F-4:	Summary of Proxies Used to Approximate Cost per Control Group Member of Similar Services	65
Exhibit F-5:	DCPD/IPEDS Variables Used in the CBA	67
Exhibit F-6:	Per-FTE Monthly Total Costs at Most-Attended Institutions	72
Exhibit F-7:	Per-FTE Monthly Total Costs at Most-Attended Institutions	72
Exhibit F-8:	Education and Training Costs Impacts by Quarter	74
Exhibit F-9:	Estimating Marginal Effective Taxes Associated with Earnings Impacts.....	76
Exhibit F-10:	Summary of Net Costs and Net Benefits by Perspective.....	78
Exhibit F-11:	Net Present Value of Quarterly Earnings after Random Assignment.....	81

Appendix A: Baseline Characteristics and Adjustments

This appendix starts with a description of the specification for baseline characteristic, including the approach to handling missing values (Section A.1). The next section compares distributions for treatment and control group members on these and other baseline measures (Section A.2), and the last section explains how the analyses control for these covariates in estimating impacts (Section A.3). It should be noted that Sections A.1 and A.2 are nearly unchanged from parallel appendices in the first, short-term report on this program (Farrell and Martinson 2017). In contrast, the approach to covariate control in Section A.3 describes some important procedural changes from those used in the prior report.

A.1 Details on Baseline Covariates

Exhibit A-1 shows the specifications and data sources for baseline covariates. Item nonresponse rates on these covariates were generally low. Across all nine PACE sites, item nonresponse rates were less than 4 percent except for parental college attendance (6.0 percent), typical high school grades (7.2 percent), family income (9.5 percent), and expected near-term future work hours (6.0 percent).

We imputed values for missing covariates using SUDAAN[®]/IMPUTE, a weighted hotdeck imputation procedure (Research Triangle Institute 2012). This imputation step entailed a single computer run on the combined sample from all nine PACE sites.¹ With this process, we replaced each missing value with an observed response from a similar case. Within specified strata, we random-matched cases with missing values to cases with reported values; we then copied over the reported value to the case where the value was missing. The strata represented a cross-classification of treatment-control status, site, National Student Clearinghouse (NSC)-reported enrollment status (*some* or *none*),² NSC-reported credential award (*some* or *none*), and number of months of NSC-reported enrollment.³

¹ Using the combined data set better controlled for school enrollment status as measured in NSC in the smaller sites.

² NSC has information on monthly enrollment and many credentials for 96 percent of college students. <https://nscresearchcenter.org/workingwithourdata/>

³ In instances where this level of matching was too restrictive because we found no matched case with a reported value, then we re-ran the procedure matching only on treatment status and NSC-reported enrollment status. In this second pass imputation, matches were allowed across sites.

Exhibit A-1: Operationalization of Baseline Measures Used as Covariates in Regression-Adjusted Impact Estimates

Variable Description	Operationalization Details	Data Source(s) (Survey Instrument: Survey Item Number)
Demographic Background		
Age	Categorical measure: Under 21 21-24 25-34 35+ ^a	BIF: B2_dob RABIT: R_RA_Date_Assigned
Female	Binary variable: 1 if female 0 if male	BIF: B7
Race/ethnicity	Categorical measure: Hispanic, any race Black, non-Hispanic White, non-Hispanic ^a Other, non-Hispanic	BIF: B9
Family structure	Categorical measure: Spouse/partner, with children Spouse/partner, without children Single, with children ^a Single, without children (Only biological and adopted children of randomized participant considered here. Stepchildren, grandchildren, younger siblings, and other children not considered.)	BIF: B13
Living with own parents	Binary variable: 1 if living with own parent(s) 0 otherwise (Presence of parents of spouse not considered.)	BIF: B13
Educational Background		
Parent attended college	Binary variable: 1 if either parent attended college 0 otherwise	BIF: B21
Usual high school grades	Categorical measure: Mostly A's Mostly B's Mostly C's or below ^a	BIF: B23
Highest level of education completed	Categorical measure: No college ^a Less than 1 year of college credit 1 or more years of college credit Associate degree or above	BIF: B17
Career Knowledge		
Career Knowledge Index (average of items)	Proportion of responses to seven questions about career orientation and knowledge to which respondent answered "strongly agree." Missing if four or more of seven responses blank.	SAQ: S13

Variable Description	Operationalization Details	Data Source(s) (Survey Instrument: Survey Item Number)
Psycho-Social Indices		
Academic discipline ^b	Average of 10 items (scale ranging 1=strongly disagree to 6=strongly agree) after reversing responses to negatively phrased items. Missing if seven or more of 10 responses blank.	SAQ: S11a
Training commitment ^c	Average of 10 items (scale ranging 1=strongly disagree to 6=strongly agree) after reversing responses to negatively phrased items. Missing if seven or more of 10 responses blank.	SAQ: S11b
Academic confidence ^d	Average of 12 items (scale ranging 1=strongly disagree to 6=strongly agree) after reversing responses to negatively phrased items. Missing if nine or more of 12 responses blank.	SAQ: S11d
Emotional stability ^e	Average of 12 items (scale ranging 1=strongly disagree to 6=strongly agree) after reversing responses to negatively phrased items. Missing if nine or more of 12 responses blank.	SAQ: S11e
Social support ^f	Average of 10 items (scale ranging 1=strongly disagree to 4=strongly agree). Missing if seven or more of 10 responses blank.	SAQ: S12
Resource Constraints (Financial)		
Family income in past 12 months	Categorical measure: Less than \$15,000 \$15,000-\$29,999 \$30,000+ ^a	BIF: B27
Received food assistance (WIC/SNAP) in past 12 months	Binary variable: 1 if yes 0 if no	BIF: B26b
Received public assistance or welfare in past 12 months	Binary variable: 1 if yes 0 if no	BIF: B26c
Financial hardship in past 12 months	Binary variable: 1 if yes to ever missed rent/mortgage payment in prior 12 months or reported generally not having enough money left at the end of the month to make ends meet over the last 12 months. 0 if otherwise	SAQ: S8, S9
Resource Constraints (Time)		
Current work hours	Categorical measure: 0-19 ^a 20-34 35+	BIF: B24
Expected work hours in next few months	Categorical measure for covariate: 0-19 ^a 20-34 35+	SAQ: S2
Expecting to attend school part-time if accepted	Binary variable: 1 if yes 0 if no	SAQ: S1

Variable Description	Operationalization Details	Data Source(s) (Survey Instrument: Survey Item Number)
Life Challenges		
Frequency of situations interfering with school, work, job search, or family responsibilities	Average of six items of frequency of problems in past 12 months (childcare, transportation, alcohol or drug use, health, family arguments, physical threats). Scale ranges from 1=never to 5=very often. Missing if four or more of six responses blank.	SAQ: S15
Stress ^a	Average of four items about feeling in control of important things and able to handle personal problems (scale 1=never to 5=very often over the past month) after reversing responses to negatively phrased items. Missing if three or more of four responses blank.	SAQ: S14

Key: BIF = Basic Information Form. RABIT = Random Assignment and Baseline Information Tool. SAQ = Self-Administered Questionnaire. SNAP = Supplemental Nutrition Assistance Program. WIC = Special Supplemental Nutrition Program for Women, Infants, and Children.

^a Category omitted in creating binary (dummy) variables for regression-adjustment models.

^b Modified version of the Academic Discipline scale in the Student Readiness Index (SRI), a proprietary product of ACT, Inc.; Le et al. (2005). Further validation in Peterson et al. (2006).

^c Modified version of Commitment to College scale in the Student Readiness Index (SRI), a proprietary product of ACT, Inc.; Le et al. (2005). Further validation in Peterson et al. (2006).

^d Modified version of the Academic Self-Confidence scale in the Student Readiness Index (SRI), a proprietary product of ACT, Inc.; Le et al. (2005). Further validation in Peterson et al. (2006).

^e Modified version of the Emotional Control scale in the Student Readiness Index (SRI), a proprietary product of ACT, Inc.; Le et al. (2005). Further validation in Peterson et al. (2006).

^f Modified version of the Social Provisions Scale; Cutrona and Russell (1987). Original scale has 24 items. This short version developed by Hoven (2012).

^g Cohen, Kamarck, and Mermelstein (1983).

A.2 Comparing Treatment and Control Groups at Baseline

Exhibit A-2 shows tests for similarity in characteristics of treatment and control group members at baseline. If the means in the two columns are congruent, then “baseline balance” was achieved. Assessment of congruence involved testing for equality of the two means separately for each characteristic.

The last column contains *p*-values for tests of hypotheses that no systematic differences exist between the treatment and control groups. If we were to repeat the randomization process a large number of times, out of 28 tests, on average, three will fall outside a 90 percent confidence interval due to chance. In this case, there were five statistically significant differences (indicated in red). We carefully reviewed data processing and other operations but could find no causes for these differences. It is likely that these are simply results due to chance. Furthermore, as described in the next section, regression adjustment helps to control for any effects that chance differences in baseline characteristics between treatment and control group members might have on the impact estimates.

Exhibit A-2: Baseline Balance

Characteristic	All Participants	Treatment Group	Control Group	p-Value
Age (%)				.252
20 or under	12.3	10.5	14.2	
21-24	20.0	19.3	20.6	
25-34	32.3	33.5	31.0	
35+	35.5	36.7	34.2	
Female (%)	83.7	83.2	84.2	.689
Race/Ethnicity (%)				.757
Hispanic, any race	46.5	46.5	46.5	
Black, non-Hispanic	21.6	22.1	21.1	
White, non-Hispanic	19.5	18.1	20.8	
Other, non-Hispanic	15.0	15.3	14.7	
Family Structure (%)				.868
Not living with spouse/partner and not living with children	46.4	46.4	46.4	
Not living with spouse/partner but living with children	28.2	29.0	27.3	
Living with spouse/partner and not living with children	12.0	11.9	12.1	
Living with spouse/partner and children	13.4	12.7	14.2	
Living with parents (%)	28.7	26.0	31.4	.058
One parent has at least some college (%)	44.4	45.7	43.1	.412
Usual High School Grades (%)				.469
Mostly A's	19.9	21.4	18.4	
Mostly B's	54.4	52.8	56.1	
Mostly C's or below	25.7	25.8	25.5	
Highest Level of Education (%)				.208
Less than a high school diploma	3.6	4.8	2.4	
High school diploma or equivalent	36.7	34.4	39.1	
Less than 1 year of college	19.4	20.0	18.7	
1 or more years of college	23.3	24.2	22.4	
Associate degree or higher	17.0	16.6	17.3	
Received vocational or technical certificate or diploma (%)	44.6	45.6	43.6	.523
Career Knowledge Index (mean)	0.56	0.60	0.52	.002
Psycho-Social Indices (means)				
Academic Discipline Index	5.55	5.59	5.51	.027
Training Commitment Index	5.73	5.74	5.71	.252
Academic Self-Confidence Index	5.11	5.14	5.08	.135
Emotional Stability Index	5.39	5.41	5.38	.305
Social Support Index	3.35	3.37	3.33	.128
Stress Index	2.13	2.10	2.16	.170
Depression Index	1.45	1.42	1.47	.118
Family Income in Past 12 Months (%)				.433
Less than \$15,000	53.1	51.0	55.2	
\$15,000-\$29,999	29.2	30.1	28.3	
\$30,000+	17.7	18.9	16.5	
Family income (mean)	\$17,319	\$17,510	\$17,124	.714

Characteristic	All Participants	Treatment Group	Control Group	p-Value
Public Assistance/Hardship Past 12 Months				
Received WIC or SNAP (%)	47.6	46.1	49.2	.346
Received public assistance or welfare (%)	19.9	21.1	18.7	.369
Reported financial hardship (%)	53.8	50.7	57.1	.056
Current Work Hours (%)				.850
0	61.9	63.3	60.6	
1-19	10.4	9.8	10.9	
20-34	16.1	15.3	17.0	
35+	11.6	11.7	11.5	
Expected Work Hours in Next Few Months (%)				.959
0	24.4	23.6	25.0	
1-19	9.2	9.5	8.9	
20-34	29.7	29.7	29.6	
35+	36.7	37.1	36.5	
Life Challenges Index (mean)	1.45	1.41	1.49	.015
Owns a car (%)	68.3	69.6	67.0	.375
Has both computer and internet at home (%)	75.1	73.8	76.4	.404
Ever arrested (%)	13.6	13.0	14.1	.616
Sample sizes	1,006	507	499	

Source: PACE Basic Information Form (BIF); Self-Administered Questionnaire (SAQ).

Note: Tests for statistically significant imbalance were based on SAS/FREQ procedure for categorical outcomes and on the SAS/TTEST procedure for other outcomes. Significant imbalances are highlighted in red, using a threshold for statistical significance of 10 percent. All values are based on baseline balance prior to imputation of missing values.

A.3 Regression Adjustment

This section describes the regression adjustment approach used to improve precision and minimize effects of sampling error on impact point estimates. In a rigorous evaluation, random assignment ensures that if the sample size is large enough, differences in average potential outcomes under treatment conditions between those assigned to the treatment group and those assigned to the control group will become vanishingly small, as will differences in average potential outcomes under control conditions between the two groups. As a result, any observed differences in average outcomes across the two groups must almost certainly be the result of treatment.⁴ Even when sample sizes are modest, random assignment ensures that that differences in average potential outcomes between the treatment and control groups arise from chance rather than biases of program operators or program evaluators. This means that unbiased estimates of the effects of treatment can be obtained by simply comparing average

⁴ Potential outcomes are a central concept in the Neyman-Rubin causal model (Holland 1986). In this model, each person has an innate pair of possible outcomes: one if treated and the other if not treated. Only one of the two potential outcomes is ever observed for each person. The average difference in potential outcomes across a specific population is said to be the local average treatment effect (LATE) or more simply, just the effect of treatment, with the context making clear the population to which it applies and supplemental analyses exploring whether the effect is homogenous within that population.

outcomes across the treatment and control groups. Moreover, it is easy to run formal tests of the hypothesis that the program has no effect (and that therefore the observed difference in mean outcomes is the result of those accidental imbalances in potential outcomes across the two groups).

Despite these favorable properties of analysis based on simple comparisons of observed means, use of regression adjustment can reduce the impact of accidental imbalances in potential outcomes across the groups, thereby increasing power to detect small program impacts (Lin 2013). To achieve this benefit, the variables used in the regression adjustment must be predictive of potential outcomes. Including other variables will increase the variance on the estimated program impact rather than decreasing it.

Opinions and practice differ on how strong the evidence for correlation between a baseline variable and the outcome must be before it makes sense to include the baseline variable in the regression adjustment.⁵ Some favor a lean approach, including just those baseline variables that have a demonstrated strong relationship to the outcome, while others favor a more comprehensive approach including all baseline variables that have a plausible theoretical relationship to outcomes of interest, believing that doing so generally bolsters confidence in study findings (Tukey 1991).

Given demands to minimize burden on participants, all measured PACE baseline variables have at least plausible relationships to PACE outcomes, but some baseline variables have been discovered to have only weak empirical relationships with PACE outcomes. Moreover, one could combine the directly measured characteristics into an unlimited number of interactions. So some judgment must be exercised about which covariates to include in regression adjustments and which to exclude.

Opinions and practice also differ on how much to customize decisions about covariate inclusion across outcomes in evaluations (like this PACE evaluation of Bridge to Employment) with multiple outcomes. A single uniform set of decisions promotes transparency, making it easier for readers to understand the procedure, while a more customized approach is likely to improve variances for at least some outcomes given that the correlation between a covariate and an outcome will vary by outcome.

In preliminary analyses for the first round of PACE reports, the team used a fairly comprehensive approach with a uniform set of decisions but discovered that this approach was causing the variances on adjusted impacts to be larger than the variances on unadjusted impacts. The discovery prompted a switch to a different approach for the first round of reports, which ultimately proved not to work as well as hoped (Judkins 2019). In response, the team developed a new approach for the current round of PACE reports. This new approach emphasizes transparency and control on imbalanced covariates, while still trying to maximize precision as far as possible given those priorities. Details follow.

⁵ For a current review of practice, see Ciolino et al. 2019.

Equation A.1 below shows the conventional regression-adjustment model:

$$Y_i = X_i\beta + \delta T_i + e_i \quad (\text{A.1})$$

where Y_i is the outcome; X_i is a row vector of baseline characteristics (hereafter referred to as covariates); β is the vector of parameters indicating the influence of each covariate on the outcome; δ is the effect of treatment; T_i is a 0/1 dummy variable indicating treatment group membership; and e_i is an error term. We fit models of this sort using SAS/SurveyReg, a procedure that uses a robust estimator of the variance of $\hat{\delta}$ and can accommodate the required nonresponse-adjustment weights for survey-measured outcomes. (See Appendix B, Section B.3 for a discussion of nonresponse-adjustment weights.)

This method is known as ordinary least squares (OLS) and has excellent properties when the sample size is many times larger than the number of baseline characteristics used as covariates (Lin 2013), even when the outcomes are not normally distributed (Judkins and Porter 2016). Estimates of the treatment effect are “asymptotically unbiased,” and under most conditions, the variance of the estimated treatment effect declines from the simple difference-in-mean-outcomes estimator of impact in proportion to the amount of outcome variation explained by the covariates.

Specifically, the relationship between the variance of the estimated treatment effect from the OLS estimation of Equation A.1 and the explanatory power of the covariates is

$\text{var}(\hat{\delta}) \approx (1 - R^2) \text{var}(\bar{Y}_T - \bar{Y}_C)$, where R^2 is proportion of the variance in Y_i explained by the baseline characteristics (X_i) in OLS estimation of Equation A.2 below:

$$Y_i = X_i\beta + e_i \quad (\text{A.2})$$

However, as mentioned above, when there are a large number of potential covariates, not all of which are useful in predicting every outcome of interest, the effect of adjustment can be the opposite of the intended effect: variances are increased rather than decreased.⁶ To avoid unnecessary variance inflation, the analyst needs to drop or otherwise reduce the influence of extraneous covariates that do not have a strong influence on the outcome of interest.

Simulation research (Judkins 2019) showed that dropping (with “backward selection”) or downweighting covariates⁷ based on simple analyses of the same data used in the evaluation

⁶ Mathematically, the presence of extraneous variables causes the coefficients of the true determinants of the outcome to be less accurately estimated. For example, if the best prediction model is $Y=2X$ but the model is fit with many extraneous covariates, the fit prediction formula could easily end up having coefficients of 1.9 or 2.1 for X instead of the best value of 2. If the wrong slope is used to correct for a treatment-control imbalance in X , the adjusted estimate of impact can easily be worse than an unadjusted estimate of impact.

⁷ An example of a method that downweights covariates is the “modified Koch method” developed for and used in the first round of PACE reports (Judkins et al. 2018; Koch et al. 1998).

yields slightly biased estimates of the variance of the estimated treatment effects (but still unbiased estimates of the treatment effect itself).⁸ This bias is negative, meaning that the variance estimates are slightly too small, making confidence intervals for impact estimates misleadingly narrow and hypothesis tests too likely to conclude that a non-zero impact has occurred when the true impact is zero or negative.

To select covariates in a manner that does not compromise variance estimation, we use the relatively recently developed technique “least absolute shrinkage and selection operator” (LASSO) with “10-fold cross-validation.”⁹ With the LASSO, the sum of absolute values of the estimated regression coefficients in Equation A.2 is constrained to be less than a preselected value (the “constraint”). If the value for this constraint is small enough, many coefficients in Equation A.2 will be forced to zero in order to fit within the cap on the sum of absolute coefficient values and thus can be removed from the list of baseline covariates. The 10-fold cross-validation is used to optimize the value of the constraint, rather than just relying on an arbitrary choice for it.

Details of the procedure are as follows:

1. With 10-fold cross-validation, the sample (both treatment and control group members) is divided into 10 equal and mutually exclusive random subsamples.
2. For each of a range of candidate values of the constraint, the LASSO procedure is run to select covariates on a sample in which one of the 10 subsamples has been dropped.
3. The model in Equation A.2 is fit on the same sample using just the variables selected in the second step for each of the candidate values of the constraint.
4. The model is used to create out-of-sample predictions of the outcome for everyone in the dropped piece of the sample, and the prediction error $\hat{Y}_i - Y_i$ is measured for each of the candidate values of the constraint.
5. Steps 2 through 4 are repeated 10 times for each candidate value of the constraint. On each iteration, a different one of the 10 subsamples is dropped. In this manner, out-of-sample prediction errors are obtained for the entire sample.
6. Mean squared prediction errors across all 10 replicates are then calculated for each of the candidate values of the constraint.

⁸ If the sample size is very large, the estimated variance on the estimated effect of treatment will be nearly unbiased even if the evaluation data are used to cull or downweight extraneous covariates. However, simulations clearly show that PACE sample sizes are not large enough to avoid biased variance estimates if “backward selection” on local data is used to prune covariates or if the modified Koch method is used to downweight extraneous covariates. Accordingly, impact analyses at three years for Bridge to Employment and all other PACE programs are not using the modified Koch method used in the first, short-term round of reports covering the first 18 months of follow-up.

⁹ See Bühlmann and van de Geer (2011) for a full explanation of these techniques.

7. The value of the constraint that minimizes this cross-validated mean squared prediction error and thus captures most of the variation reduction possible with the available covariates is selected as the optimal constraint.¹⁰ Whichever variables have nonzero coefficients in the model for that optimal constraint are used as covariates in the impact regressions. All other baseline characteristics are discarded. All of this is done automatically in SAS[®]/GLMSELECT. Simulations under PACE-like conditions (in terms of sample sizes and the numbers of covariates) when developing the analysis plan for the entire suite of PACE three-year reports (Judkins et al. 2018) demonstrate that this technique reduces the true variances without biasing variance estimates.¹¹

In principle, we could repeat the LASSO with 10-fold cross-validation independently for every outcome for each of the nine PACE programs. But such an approach would produce a different final covariate list for each outcome and program, leading to some loss in transparency and making it harder for outside researchers to replicate the PACE results. At the other extreme, we could run the LASSO just once for each program for the most important confirmatory outcome and then use the resulting set of selected covariates for all impact estimates for the program. But we believe that this would result in more precision loss than can be justified for the sake of transparency.

As a compromise between these extremes, we selected one set of covariates for each of three domains and customized them for each of the nine PACE programs. The three domains are (1) analyses of *employment and earnings* outcomes that are conducted on the dataset of merged data from the three-year follow-up survey and the National Directory of New Hires (NDNH); (2) analyses of *education* outcomes (whether based on the survey, NSC, or local or state college records); and (3) analyses of all *other* outcomes (most of which concern personal and family well-being and economic independence). The pool of potential covariates was the same for all three domains—with one important exception: indicators of pre-baseline earnings based on NDNH data are only allowed in analyses of NDNH-based outcomes.¹²

To identify covariates for this report, we ran the LASSO procedure for the most salient outcome within each of the three domains (*earnings and employment*, *educational progress*, *other*) at each of the nine PACE programs.¹³ For NDNH analyses, the confirmatory outcome is average

¹⁰ One could simply use the LASSO to select covariates with a pre-specified value of the constraint, but the 10-fold cross-validation provides a principled method for selecting the constraint.

¹¹ See Judkins (forthcoming) for additional detail.

¹² This is because we analyzed survey outcomes on Abt's secure server rather than on the ACF secure server. Though both systems have very high security procedures, agreements with the Office of Child Support Enforcement (OCSE) permit the NDNH data to reside only on the ACF secure server. It would have been possible to analyze all survey outcomes on the ACF secure server, but doing so would have significantly burdened the study's analytic operations without any commensurate benefit. It would also prevent us from analyzing survey data for people whose names and Social Security numbers do not properly match the Social Security Administration's records.

¹³ Selection started with the set of baseline covariates used in the analyses of follow-up data at 18 months after random assignment for the short-term impact report (shown in Exhibit A-3).

quarterly earnings for the 12th and 13th quarters after randomization (Q12, Q13), so that is a natural choice for the outcome around which to optimize covariate selection. In the educational progress domain, the most important outcome varies by PACE program. As discussed in the body of this report, for Bridge to Employment in the Healthcare Industry, the most salient education outcome is any healthcare credential from a school. As the most salient outcome for the third domain, we selected whether anyone in the household draws means-tested public benefits. We made this last decision because of the centrality of the concept of self-sufficiency in the rationale for creating the PACE project.¹⁴ We made these choices prior to reviewing any impact estimates.

In addition to covariates based on the above procedures, regression models included covariates for which baseline distributions differ for treatment and control group members at the 5 percent level.¹⁵

Exhibit A-3 below shows the covariates that we selected with the LASSO procedure or by virtue of their being out of balance (OOB) at baseline. Some covariates were selected both by the LASSO and by virtue of being out of balance at baseline. When this happened, the covariate is so flagged in the table. For categorical variables, the LASSO procedure worked on dummy variables for the individual levels; so for a variable with four levels, it was possible for just one of three dummy variables to be selected. In contrast, the out-of-balance test selected all or none of the levels of a categorical variable. The table shows all possible levels of categorical variables and indicates which specific categories we selected as covariates.

¹⁴ The original name for PACE was “Innovative Strategies for Increasing Self-Sufficiency.” The promotion of self-sufficiency is also central to the goals of the career pathways framework as articulated in Fein 2012.

¹⁵ Baseline balance was assessed prior to imputation of missing data. See Exhibit A-2.

Exhibit A-3: Covariates Selected, by Outcome Domain

Baseline Covariate	Domain		
	NDNH-Based Employment and Earnings	Educational Progress	Other
Age 20 or under 21-24 25-34 35+			
Gender Female Male			LASSO
Race/Ethnicity Hispanic, any race Black, non-Hispanic White, non-Hispanic Other, non-Hispanic			LASSO
Family Structure Not living with spouse/partner and not living with children Not living with spouse/partner but living with children Living with spouse/partner and not living with children Living with spouse/partner and children		LASSO	LASSO
Living with parents			
One parent has at least some college			
High School Grades Mostly A's Mostly B's Mostly C's or below		LASSO	
Current Education High school equivalent or less Less than 1 year of college 1 or more years of college Associate degree or higher			
Career Knowledge Index	OOB	OOB	OOB
Family Income in Past 12 Months Less than \$15,000 \$15,000-\$29,999 \$30,000+			LASSO LASSO
Pre-Randomization Quarterly Earnings (NDNH) 4 quarters prior to randomization 3 quarters prior to randomization 2 quarters prior to randomization 1 quarter prior to randomization	LASSO	Not available	Not available
Psycho-Social Indices Academic Discipline Index Training Commitment Index	OOB	OOB	OOB

Baseline Covariate	Domain		
	NDNH-Based Employment and Earnings	Educational Progress	Other
Academic Self-Confidence Index			
Emotional Stability Index			
Stress Index			
Life Challenges Index	OOB	OOB	OOB
Public Assistance/Hardship Past 12 Months			
Received WIC or SNAP			LASSO
Received public assistance or welfare			LASSO
Reported financial hardship	OOB	LASSO, OOB	OOB
Current Work Hours			
0-19			
20-34			
35+			
Expected Work Hours in Next Few Months			
0-19			
20-34			
35+		LASSO	
Plan to attend school only part-time if admitted to Bridge to Employment		LASSO	

Key: SNAP = Supplemental Nutrition Assistance Program. WIC = Special Supplemental Nutrition Program for Women, Infants, and Children.

Note: "LASSO" flags that the covariate was selected by the LASSO for variance reduction. "OOB" flags that the covariate was selected because it was significantly out of balance.

Exhibit A-4 below shows impacts on selected confirmatory and secondary outcomes before and after regression adjustment without weights.¹⁶ Though the two sets of estimates lead to similar conclusions, the estimates of program impacts appear to be more sensitive to the choice to use regression adjustment than are the standard errors. Contrary to the primary goal of regression adjustment, less than a half of the regression-adjusted standard errors are smaller than the unadjusted standard errors.¹⁷ This disappointing result is probably due to the decision to only weakly customize covariate selection to outcomes and to place a greater focus on choosing imbalanced covariates than on choosing predictive covariates.

The impact of adjusting for out-of-balance but only marginally relevant covariates is most evident in the estimate of program impact on confidence in career knowledge. The baseline scores for this variable were not predictive of *educational progress* or *economic well-being* at three years (not shown), but they are strongly predictive of *career knowledge* at three years

¹⁶ We did not use the weights in the preparation of this table because they are not required for the first panel (Full Sample), and because in this section we want the focus to be on the role of covariates. See Appendix Exhibit B-11 for the impact of nonresponse-adjustment weights on these estimates.

¹⁷ The prominent exception is receipt of public benefits. The standard error on the estimated impact of the program on this outcome is 7 percent smaller after adjustment (3.91 versus 3.67). This is probably because this outcome was chosen as the critical outcome among all other life outcomes.

(also not shown) and they were strongly imbalanced at baseline (the score of .60 on Career Knowledge Index in the treatment group was .08 higher than the score of .52 on the control sample; Exhibit A-2). As a result, the effect of covariate adjustment on the estimated impact of the Bridge to Employment program on career knowledge is also strong. As seen in Exhibit A-4, without regression adjustment, the estimated program effect of 0.069 is statistically significant at the 10 percent level even though it is smaller than the initial imbalance. With regression adjustment, the estimated impact is near zero (0.017). This smaller impact estimate is probably closer to the truth given the initial imbalance, so this is an example where controlling on out-of-balance covariates can be advantageous even when the covariates are only marginally relevant to the more important outcomes.

Exhibit A-4: Comparison of Confirmatory and Secondary Impact Estimates Unadjusted and Adjusted for Baseline Imbalances

Domain (Data Source), Outcome	Unadjusted Estimate		Adjusted Estimate	
	Impact	Standard Error	Impact	Standard Error
Confirmatory Outcome (NDNH)				
Full Sample				
Average quarterly earnings Q12-Q13 after randomization (\$)	289	276	297	254
Secondary Outcome: Employment (Survey)				
Survey Respondents without Weights				
Employed at survey follow-up (%)	6.7**	3.6	6.2**	3.6
Employed at \$15 per hour or above (%)	2.8	3.5	0.6	3.5
Employed in a job requiring at least mid-level skills (%)	-1.6	3.4	-3.0	3.4
Employed in the healthcare field (self-classification)	12.4***	3.8	11.3***	3.9
Secondary Outcome: Education (Survey)				
Survey Respondents without Weights				
Full-time-equivalent months enrolled at any school (months)	-0.4	0.7	-0.7	0.7
Receipt of a healthcare credential from any school (%)	17.5***	3.8	16.9***	3.8
Receipt of an exam-based certification or license (%) ^a	6.4**	2.9	6.9**	2.9
Secondary Outcome: Other (Survey)				
Survey Respondents without Weights				
Indicators of Independence and Well-Being				
Has health insurance coverage (%)	3.6*	2.7	2.6	2.7
Receives means-tested public benefits (%)	-2.8	3.9	-0.3	3.7
Student debt (\$)	-1,400	749	-1,650	760
Financial hardship (%)	-2.9	3.9	0.9	3.8
Indices of Self-Assessed Career Progress (average)				
Confidence in career knowledge ^b	0.07*	0.05	0.02	0.04
Access to career supports ^c	-0.03	0.02	-0.02	0.02
Sample sizes (across treatment and control groups):				
NDNH	1,004			
Survey	658			

Source: PACE 18-month follow-up survey, PACE three-year follow-up survey, National Directory of New Hires.

^a Blended 18-month and three-year survey results.

^b Seven-item scale tapping self-assessed career knowledge; response categories range from 1=strongly disagree to 4=strongly agree.

^c Six-item scale tapping self-assessed access to career supports; response categories range from 1=no to 2=yes.

Statistical significance levels, based on one-tailed *t*-tests of differences between research groups: *** 1 percent level; ** 5 percent level; * 10 percent level.

Appendix B: Three-Year Survey Data

This appendix documents key technical detail underlying analyses of the three-year follow-up survey data.¹⁸ Section B.1 documents coding for scales based on follow-up survey data. Section B.2 describes the imputation process for some missing survey data elements in the construction of outcomes. Section B.3 analyzes survey nonresponse and documents the process we used to build the nonresponse weights used in the impact analysis. Sections B.4 and B.5 present evidence about the quality and completeness of survey responses. Before getting into those details, we provide an overview of the measurement goals and structure of the instrument.

The survey sought to collect a complete history of jobs and periods of schooling since randomization (including the progression and interleaving of these spells), credits and credentials earned; earnings growth, and self-employment. In addition, the survey measured several psycho-social skills, family formation and growth, income and material well-being, and child outcomes.

The Integrated Training and Employment History module of the three-year survey aimed to collect a complete history of training and employment between randomization and the day of interview some three years later. Given data collection plans, the approach needed to work over the phone. The instrument development team reviewed several past efforts to collect such histories, but only one of the past approaches seemed likely to be workable over the phone¹⁹—an approach developed for a German survey instrument that studies the training and work histories of German youth.²⁰ To our knowledge, this was the first time that the German approach had been attempted in the United States.

Conceptually, a history could be built either forward from randomization or backward from the day of interview. The German study worked forward with apparent success, so the team also adopted that approach. One modification we made was to take each respondent through his or her training and employment history twice instead of just once. First, the survey collects the spell history (dates, whether work or school, and place names). This is the “scaffolding.” Once the scaffolding has been built, the interviewer takes the respondent back through the history a

¹⁸ The full instrument is available at <http://www.career-pathways.org/career-pathways-pace-three-year-instrument/>.

¹⁹ The other approaches involved the use of paper calendars that could be spread on the table and jointly viewed. Clearly this would not be feasible over the phone.

²⁰ The 2011 BIBB Transitional Study was a retrospective longitudinal survey conducted by the *Bundesinstitut für Berufsbildung* (Federal Institute for Vocational Education and Training) on a representative basis that recorded in detail the whole of the educational and occupational biographies of persons born between 1987 and 1992 and resident in Germany. For full details: Beicht and Friedrich (2008). For a brief English synopsis of one report using some of the survey data: <http://www.bibb.de/en/64317.htm>

second time to systematically collect more information about each training spell. There are two motivations for this two-pass approach.

1. By asking the respondent to focus on one type of information at a time, collection of date data may be more consistent across spells.
2. This approach allows for more straightforward programming.

B.1 Measures Based on Follow-up Survey Data

Exhibits in this section detail the operationalization of survey-based outcomes used in impact analyses in the main report. These exhibits also reference the underlying survey questions. B-1 provides details on outcomes in the education domain, as reported in Chapter 3. Exhibits B-2 and B-3 provide similar details on outcomes in the employment/earnings and intermediate outcomes, respectively, as reported in Chapter 4. Finally, Exhibits B-4 and B-5 do the same for other life outcomes domains and child outcomes, respectively, as reported in Chapter 5.

The measurement of earned credentials involved use of many questions from the follow-up survey. Rather than repeat that information for each outcome, this section first provides a review of the strategy for this task. The survey asked about credentials in three different ways. The survey first asked whether they had received “a diploma, certificate, or academic degree for completing any regular college classes.” Among those answering the question affirmatively, the survey asked for a list of such credentials and more about each one, including issuing school, award date and (for sub-degree credentials) the name and typical length of study required to earn it. The survey then asked whether they had received “any diplomas or certificates from a school for completing any vocational training.” Among those answering the question affirmatively, the survey asked for a list of such credentials with no further detail beyond issuing school and award date. Finally, the survey asked whether they had “received a professional, state, or industry certification, license, or credential from an authority other than a school.” Among those answering the question affirmatively, the survey asked for a list of such credentials, award dates, and the type of authority issuing it. In post-survey processing, PACE evaluation staff imputed the required length of study required to earn the credential based on the respondent-provided name of the credential. This imputation had three levels: less than a year, a year or more but not as much as for an associate degree, as much as an associate degree, or as much as a bachelor’s degree.

Exhibit B-1: Details on Specifications for Survey-Based Education Outcomes in Chapter 3

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Secondary Outcomes		
Education		
Full-time-equivalent months enrolled at any school	Students were asked for the dates of attendance of each school attended and their status while enrolled. If their status was "part-time," then the number of months was multiplied by 0.25 to estimate full-time-equivalent months. Similarly, if their status was "equal mix," then number of months was multiplied by 0.50 to estimate full-time-equivalent months. We developed this rule based loosely on guidance in NSC documents about how schools should classify less-than-full-time enrollment. Because the survey response categories were different from those used in the NSC and because students might have different understandings than schools did, this decision was fairly arbitrary. Alternate rules might have worked just as well.	C2, C3, D2
Received a healthcare credential from any school	Respondents were asked whether they had received "a diploma, certificate, or academic degree for completing any regular college classes" and whether they had received "any diplomas or certificates from a school for completing any vocational training." Both types of credentials were listed by name, and respondents were asked for each one whether the credential was "related to working in the field of healthcare."	I2, I2b, I2c, I3, I3b, I3c
Received an exam-based certification or license	Respondents were asked whether they had "received a professional, state, or industry certification, license, or credential from an authority other than a school." This measure uses the 18-month survey for exam-based credentials reported through the time that survey was completed and uses the three-year survey for exam-based credentials that were reported to be earned after completion of the short-term survey.	Three-year: I3d, I3di, I3h 18-month: A56, A56a
Exploratory Outcomes		
Full-time-equivalent months enrolled at a college	School names were matched to IPEDS. If IPEDS indicated that the school was degree granting, then the school was considered a college.	C1, C2, C3, D2
Full-time-equivalent months enrolled at another school	School names were matched to IPEDS. If IPEDS indicated that the school was not degree granting, then the school was considered another school. Schools not matched to IPEDS were considered non-Title IV provided that we could locate a website for the school.	C1, C2, C3, D2
Received any type of credential from any school	Any credential from a school. Excludes credentials issued by other authorities.	I2, I2a_1, I2b, I2c, I3
Received a healthcare credential from a college	Similar to received a healthcare credential from any school, but restricted to schools that are accredited to issue degrees, as documented in the IPEDS database.	C1, I2, I2a_1, I2b, I2c, I3
Received a healthcare credential from another school	Credentials in the field of healthcare that were awarded by schools other than colleges. The school could be listed in IPEDS database as a school that does not issue degrees (Carnegie level 3) or it could be unlisted in IPEDS and therefore a non-title-IV school. Almost all such schools are private, and most are for profit.	C1, I2, I2a_1, I2b, I2c, I3
Received any type of credential from a college	Credentials awarded for taking regular college classes or for vocational classes at a college.	C1, I2, I2a_1, I2b, I2c, I3
Received any type of credential from another school	Credentials awarded by schools other than colleges.	C1, I2, I2a_1, I2b, I2c, I3

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Enrolled in training or education at survey follow-up	Determined based on reported dates of enrollment in education and training activities and date of interview.	Most of modules B, C, and E

Key: IPEDS = Integrated Postsecondary Education Data System. NSC = National Student Clearinghouse.

Exhibit B-2: Details on Specifications for Survey-Based Employment/Earnings Outcomes in Chapter 4

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Secondary Outcomes		
Employed at survey follow-up	Determined based on reported dates of jobs and date of interview.	Most of modules B, C, and E
Career Progress		
Employed at \$15 per hour or above	Analyzed response to survey question for control group. Selected \$15 per hour as the threshold because it was close to the 60th percentile of hourly wages among employed control group members. This percentile was picked as being a reasonable goal for graduates of Bridge to Employment.	F5
Employment in job requiring at least mid-level skills	Three open-ended questions about the kind of work done, the usual activities completed, and the job title were coded into an SOC code. We then looked up the Job Zone ^a for each SOC code in the O*NET system. ^b Job Zone 3—occupations that need medium preparation—seemed a reasonable goal for graduates of Bridge to Employment.	G2a, G3, G4
Employment in the healthcare field (self-classification)	Currently employed respondents were asked whether they were “employed in a healthcare job.”	G2
Exploratory Outcomes		
Works at least 32 hours per week	Currently employed respondents were asked about their typical hours worked.	F6
Currently employed, working straight day, evening, or night shifts	Currently employed respondents were asked about their typical work schedule. Answer possibilities included straight shifts, rotating shifts, split shifts, irregular schedules, and other.	G6, G6a
Currently working in a job that offers health insurance	Currently employed respondents were asked whether health insurance was available through the employer as a fringe benefit.	G8a
Currently working in a job with a supportive working environment	Questions about job benefits and conditions were used to cluster jobs into three categories. Jobs in this category generally provided employees with flexibility to balance work and family, a supportive set of co-workers and supervisors, a rich set of benefits, and opportunities for advancement.	G7, G8a-G8e, G9, G10
Working in a healthcare occupation	Three open-ended questions about the kind of work done, usual activities completed, and the job title were coded into a SOC code. If the first two digits of the SOC were 29 (Healthcare Practitioners and Technical Occupations) or 31 (Healthcare Support Occupations), then the respondent was considered working in a healthcare occupation. ^c	G2a, G3, G4

Key: SOC = U.S. Department of Labor Standard Occupational Classification.

^a <https://www.onetonline.org/help/online/zones> [accessed September 12, 2016].

^b <https://www.onetonline.org/> [last accessed September 12, 2016]. There are five Job Zones. A Job Zone is a group of occupations that are similar in education needed to do the work, related experience needed to do the work, and amount of on-the-job training needed to do the work. Job Zone 3 is described in the O*NET system documentation as “Employees in these occupations usually need one or two years of training involving both on-the-job experience and informal training with experienced workers. A recognized apprenticeship program may be associated with these occupations.”

^c Being employed in a healthcare occupation is usually associated with employment in the healthcare industry, but this is not always true. School nurses are one example of a healthcare worker being employed in an industry other than healthcare. Conversely, many people employed in the healthcare industry are not healthcare workers. Hospital janitors are one example. The survey did not ask about industry of employer.

Exhibit B-3: Details on Specifications for Survey-Based Intermediate Outcomes in Chapter 4

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Secondary Outcomes		
Access to career supports	<p>This was a new scale created for PACE at 18-month follow-up. It is a six-item scale counting number of types of career-supportive relationships in workforce and education settings. The motivation for creating this scale was the theory that richer social networks are one of the benefits of higher education (e.g., Goldrick-Rab and Sorensen 2010).</p> <p>Say you need advice of help in taking a next step on a career pathway of interest to you. Please tell me if there is anyone you'd be comfortable turning to:</p> <ul style="list-style-type: none"> • Who has a college degree? • Who is currently going to college? • Who works at a local college, either as a teacher or staff member providing help to applicants or students? • Who works for a local community organization helping people find education and training, work, and related supports? • Who works in an occupation of interest to you? • Who has a management job in a work setting matching your career interests? 	K4
Confidence in career knowledge	<p>This seven-item scale was based on a review of six survey instruments as well as literature. The first two scale items (a, b) were adapted from the Career Decision Self-Efficacy–Short Form (Betz and Taylor 2001). Three items (d-f) were adapted from the Career Exploration Survey (Stumpf et al. 1983). Two items (c, g) were new and written specifically for the PACE Basic Information Form. Response categories ranged from 1=strongly disagree to 4=strongly agree.</p> <ul style="list-style-type: none"> a. You know how to accurately assess your abilities and challenges? b. You know how to make a plan that will help achieve your goals for the next five years? c. You know how to get help from staff and teachers with any issues that might arise at school? d. You know the type of job that is best for you? e. You know the type of organization you want to work for? f. You know the occupation you want to enter? g. You know the kind of education and training program that is best for you? 	K6
Exploratory Outcomes		
Perceived career progress	<p>This was a new scale created for PACE at 18-month follow-up. It is a three-item scale of self-assessed career progress. Response categories range from 1=strongly disagree to 4=strongly agree. It was designed specifically to measure a respondent's sense of progress in a career pathways program as described in Fein (2012).</p> <ul style="list-style-type: none"> • I am making progress towards my long range educational goals • I am making progress towards my long-range employment goals • I see myself on a career path 	I5, I6

Exhibit B-4: Details on Specifications for Survey-Based Other Life Outcomes in Chapter 5

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Secondary Outcomes		
Student debt	Students were asked about personal borrowing to go to school since randomization. For those who had difficulty answering the question about the exact amount, a categorical response option was offered. These were then imputed to continuous levels.	M6, M6a
Has health insurance coverage	Includes the offer of healthcare by employer or actual receipt if not offered by employer.	G8a, M12
Receives means-tested public benefits	Respondents were asked whether they or anyone else in their household received TANF, SNAP, WIC, Medicaid, subsidized childcare, Section 8 or Public Housing, LIHEAP, or FRPL.	M3a, M3b, M3c, M3e, M3f, M3g, M3h, M3i
Any signs of financial distress	For the three-year follow-up, this scale is an expanded version of the financial hardship measure used in 18-month follow-up. It flags any signs of financial distress in terms of troubles paying bills (rent/mortgage, gas/oil/electricity), utility disconnects (gas/electric/oil or telephone), delayed healthcare, delayed dental care, delayed prescription drug procurement, not having enough to eat (sometimes or often), or not having enough money to make ends meet at the end of the month.	M9a-g, M10, M11
Exploratory Outcomes		
Unsecured debt of \$5,000 or more	Respondents were asked about debt other than student debt and secured debt (such as mortgages or title loans). Debts in the name of their spouse or partner were included.	M8
EITC claimant	Respondents were asked about actual or intended claim of the Earned Income Tax Credit for the prior tax year.	M5
Dependent on family	If the respondent reported cash transfers in the prior month from friends or family outside the home or living rent-free with friends or family.	M1k, M13
Homeowner	Reports current living situation is "own your own home."	M13
Homeless	Reports that current living situation is group shelter or gave an open-ended answer that indicated respondent lived in a place not meant for human habitation or a homeless shelter, or couch surfed (slept on couches at various friends' houses).	M13
Parental student debt	Respondents were asked about borrowing by parents on behalf of the student to go to school since randomization. For those who had difficulty answering the question about the exact amount, a categorical response option was offered. These were then imputed to continuous levels.	M7, M7a

Key: EITC = Earned Income Tax Credit. FRPL = free or reduced-price lunch. LIHEAP = Low Income Home Energy Assistance Program. SNAP = Supplemental Nutrition Assistance Program. TANF = Temporary Assistance for Needy Families. WIC = Special Supplemental Nutrition Program for Women, Infants, and Children.

Exhibit B-5: Details on Specifications for Survey-Based Child Outcomes in Chapter 5

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Exploratory Outcomes		
Children of All Ages		
Parent believes child will graduate college	Parent asked how far child will go in school. Outcome equals 1 if parent reports child will finish college or if parent reports child will earn advanced degree after college; 0 otherwise.	P1
Highly engaged parent	<p>This is a new scale developed for the three-year evaluations of PACE and HPOG 1.0. It was based on imputed average hours of time per day spent with the child in the typical week. The algorithm was different for preschoolers versus school-age children. Both thresholds were set at the 75th percentile for all children in the pooled evaluation samples for PACE and HPOG 1.0.^a</p> <p>For preschoolers, parents were credited with 1 hour for each shared breakfast in the typical week; 1 hour for each shared dinner; 7 hours if they usually put the child to bed; and 1.5 hours if they read to the child once or twice a week, 4.5 hours if they read to the child three to six times a week, and 7 hours if they read to the child every day. These hours were summed and then divided by 7. The maximum value was 4 and the 75th percentile was 3.64. If the quotient was greater than this percentile the parent was said to be highly engaged with the preschooler.</p> <p>For school-age children, parents were credited with 1 hour for each shared breakfast in the typical week, 1 hour for each shared dinner, 7 hours if they usually put the child to bed, 7 hours if they were usually present before the child leaves for school, 7 hours if they were usually present after the child comes home from school, 7 hours if they were usually present after dinner, and 7 hours if they were present with the child during the weekend. These hours were summed and then divided by 7. The maximum value was 8 and the 75th percentile was 7.28. If the quotient was greater than this percentile, the parent was said to be highly engaged with the school-age child.</p>	O3a, O4a, O5a, O6a, O7a, O7b, O7c, P3, P6
Parent self-efficacy for helping child navigate school	Existing scale. ^a The seven-item scale captures parents' beliefs about their capability to help their child succeed in school. Response categories ranged from 1=disagree very strongly to 6=agree very strongly.	P9
School-Age Children		
Child repeated any grades	Yes/no question if child repeated any grades in school.	Q10
Days child late for school last month	How many days was child late for school in last month (if in summer vacation, asked about last month child was enrolled in school).	Q12
Days child absent from school last month	How many days was child absent from school in last month (if in summer vacation, asked about last month child was enrolled in school).	Q11

^a ACF's Health Profession Opportunity Grants (HPOG) Program, like PACE, provides training to low-income individuals, but specifically for healthcare occupations. A first round of grants was awarded in 2010 (HPOG 1.0). Three of the nine programs studied in PACE were HPOG 1.0 grantees. For more: <https://www.acf.hhs.gov/ofa/programs/hpog>.

^b Walker et al. (2005).

B.2 Imputation in the Three-Year Survey

As in any survey, some respondents do not answer every question. Rather than dropping respondents with missing survey items, we used a variety of approaches to make use of the partial responses. Our decision to include or drop such cases varied, depending on whether the unanswered question was embedded in a sequence in which all questions needed to be answered to calculate the value of a scale, whether the unanswered question was embedded in a block of unanswered questions, and the frequency of nonresponse to the question across respondents.

The default rule was to drop persons from any analysis involving unanswered question but to include them for all other analyses. Where this rule would result in a sharp drop in sample size—either for the question by itself or for a scale involving the question—then we instead imputed responses for those people for those questions, rather than dropping them. Additionally, we imputed blocks of responses for two groups of people: those with large blocks of missing data and those who, based on administrative data, appeared to have failed to report one or more education spells.

The goals of imputation were variance and bias reduction. Both goals are achievable with the rich set of parallel outcomes measured in the three-year survey. For example, indications of problems paying bills is valuable information for imputing missing income. We made decisions for all PACE sites on a global basis. Either we implemented an imputation procedure for a question in all nine PACE sites, or we left the question blank in all sites. Specifically, we imputed seven types of missing data:

1. Number of college credits;
2. Credential award dates;
3. Income (personal and household);
4. Early certificates and licenses (first 18 months after randomization);
5. Skipout (i.e., missing data on spells caused by trying to avoid respondents ending the survey);
6. Spell start and end dates (job spells and school spells); and
7. Survey data on school spells reported to the National Student Clearinghouse (NSC) but not by respondent.

This section briefly describes each of these imputations and their prevalence. We used a common methodology for the first four types of missing data. Section B.2.1 provides the detail on these imputations. Section B.2.2 gives details on the imputation methodology for the other three types of missing data.

Types and Rates of Imputation. Exhibit B-6 below lists the seven types of imputation and shows the imputation rates for the survey respondents in the evaluation sample for Bridge to Employment. The instrument asked about credits spell by spell. It was fairly common for

respondents to be unable to recall the number of credits they had earned during one or more training spells. They also had trouble recalling the dates on which they received credentials. Income was also frequently missing. The instrument prompted respondents to give a categorical answer (“bracketing”) if they could not give an exact figure.

Exhibit B-6: Imputation Rates among Survey Respondents in the Evaluation Sample for Bridge to Employment

Type of Imputation	Job Spells (%)	School Spells (%)	Credentials (%)	People (%)
1. Number of college credits	n/a	n/a	n/a	11.8
2. Credential award dates	n/a	n/a	8.5	n/a
3. Income				
Personal (categorical)	n/a	n/a	n/a	5.2
Personal (exact)	n/a	n/a	n/a	10.0
Household (categorical)	n/a	n/a	n/a	15.5
Household (exact)	n/a	n/a	n/a	26.1
4. Early certifications and licenses	n/a	n/a	n/a	14.1
5. Skipout	8.3	9.5	10.2	8.5
6. Spell start and/or end dates (job, school)	8.7	13.0	n/a	n/a
7. Survey data on school spells reported to NSC but not by respondent	n/a	7.8	5.8	4.7

Source: PACE three-year follow-up survey.

Note: Exact income was missing more often than categorical income because respondents unable or unwilling to provide an exact amount were encouraged to report a bracketed amount. n/a indicates not applicable.

The “Early Certifications and Licenses” row refers to our decision to impute this outcome for the 18-month follow-up survey for those study participants who were not interviewed at 18 months after randomization but who were interviewed at three years. This imputation involved creating a composite scale using the 18-month interview to measure receipt in the first 18 months and the second interview to measure receipt in the second 18 months. Section B.4 provides information about the rationale for this composite scale.

The “Skipout” row refers to block missingness in the Integrated Training and Employment History module. The German survey upon which this module was modeled experienced a high level of break-off (12 percent), meaning people discontinued the interview midstream and declined to restart it. To prevent similar problems for this three-year analysis, the PACE survey added a skipout feature in the module. If a person refused to answer any question in the module or gave a response of “don’t know” to any of several critical flow-controlling questions in the module, the interview flow automatically skipped ahead to the next modules (e.g., on 21st century skills, family structure, income and material well-being, and child outcomes).²¹ With this approach, complete interview breakoffs were nearly eliminated, but a large block of missing

²¹ The original intent was not to skip past questions about credential attainment and current job conditions, but a mistake in the specifications caused these sections to also be skipped.

data was created for about 7 percent of respondents (across the entire three-year sample)—much lower than the break-off rate on the German study, but still high enough to require special attention.

Nonresponse was non-negligible for start and end dates of both job and school spells, particularly start dates. This is not surprising given that the reference period was up to three years long (and longer for people interviewed later in the survey period and for spells that started prior to randomization).

The final row of Exhibit B-6 refers to an adjustment for undercoverage of NSC-reported spells. This adjustment started with a match of survey reports with administrative data on college attendance from the NSC. We flagged respondents who had spells of college attendance according to the NSC but who did not themselves report any training (college or other type of school) since randomization. Although the NSC is not error-free, its enrollment coverage is generally high. Accordingly, we imputed all the data from the matched NSC spells to survey respondents who did not report such spells.

B.2.1 College Credits, Credential Award Dates, Income, and Early Certificates and Licenses (Imputations 1-4)

As mentioned above, four of the seven types of imputation utilized a common imputation procedure: college credits, credential award dates, income, and certifications and licenses in the first 18 months. This section discusses the basic procedures used and provides additional details for each of the four types of missing data.

Core Imputation Procedure. The core imputation methodology involved a number of steps. The first step entailed assembling a list of potential predictors and imputing any missing data in them.²² The list of potential predictors included program, treatment status, the interaction of program with treatment status, baseline variables, parallel outcomes, and two-way and three-way interactions of both baseline variables and parallel outcomes with program and treatment status.

The second step entailed the use of a cross-validated LASSO procedure to fit a linear model for the target variable in terms of the assembled predictor list.²³ We did this on a pooled dataset that contained respondents from all nine PACE sites ($n=6,773$, of whom 5,910 responded to both follow-up surveys) and sometimes respondents from HPOG-only programs, as well. Note that though this procedure allowed program, treatment, their interaction with each other, and their interactions with many other variables to enter the model, it did not force any of them in. We discuss the implications of this decision after first finishing a description of the procedure.

²² The only purpose of this imputation was to facilitate automated variable selection in the next step. After we used these imputed values in the predictors to predict new exam-based certifications and licenses as of the time of the 18-month survey, we discarded them. We carried out this imputation with SAS/MI/FCS.

²³ See Appendix A.3 for details on the cross-validated LASSO.

The third step used predicted values from the final linear model to create a nested set of three partitions for each combination of site and treatment status.²⁴ The finest partition involved splitting the sample into 20 equal-sized groups based on the predicted probability of having reported an exam-based certification or license if respondents had been interviewed at 18 months. The middle partition corresponded to deciles of this same probability, and the coarsest partition corresponded to quintiles of this same probability.

The fourth and final step used the hotdeck imputation procedure in SUDAAN to randomly match each nonrespondent with a respondent within cells defined by PACE program, PACE treatment status, and the nested partitions. Most cases were matched within cells defined by the 20-level partition. If there were no matches within those cells, then the procedure sought matches within the coarser partitions, first with the 10-level version and then with the five-level version if necessary. If even that did not permit a match, then the procedure randomly matched any unmatched nonrespondents with any respondent in the same PACE program with the same treatment status.

We ran the final hotdeck procedure five times with different random seeds to produce multiple imputations. We used these multiple imputations in the formal analysis runs to add between-imputation onto the naïve variance estimates on the full sample, using Rubin's classic formula.²⁵

We now return to the implications of our decision not to force the interactions of site and treatment group with every other variable in the model. First, it is critical to note that we constrained matches to be from the same site and treatment group. This provided strong protection against imputation-caused bias in the estimated treatment impact. We used the models from the pooled dataset only to guide the matching of respondents and nonrespondents *with the same treatment status in the same site*. One way to think of this is that we used the pooled dataset to define a distance metric that we then applied within site and treatment group. An alternative procedure would have been to just randomly match respondents and nonrespondents within cells defined by site and treatment group. The point of using a distance metric rather than randomly matching is to reduce variance and the possibility of nonresponse bias. For a site with a large sample size, forcing in all the interactions of site and treatment group with other variables might not cause much deterioration in model quality, but in small sites forcing would almost certainly have made it more difficult to detect subtle main effects.²⁶

Life Trajectory Clusters. The survey contained multiple measures of financial and social-emotional well-being. We theorized that these variables would be useful predictors of several types of missing data, particularly the missing data created by skipouts because none of these questions were involved in the bad skip pattern. However, interpretation of high-dimensional models is difficult. As a way of incorporating these rich data on well-being into imputation

²⁴ A "partition" of a sample is an exhaustive and mutually exclusive collection of subsets of the sample.

²⁵ See for example, Rubin (1987).

²⁶ Algorithmically, the way to force in all interactions is to run the LASSO on a dataset restricted to just the cases in a particular site and treatment group. Even for the largest PACE site, this would not have provided nearly as much power to detect subtle main effects.

models while still keeping the models fairly easy to interpret, we condensed all these measures into a partition of the sample using cluster analysis. We were able to identify five clusters of respondents who vary clearly in terms of quality of life and core self-evaluation and family dependence. For shorthand, we refer to them as “life trajectory” clusters because one of the variables that they vary on most clearly is a sense of career progress:

- “Overextended”—above average income but also above average financial stress and low scores on psycho-social skills.
- “Family supported”—below average income but strong family supports that protect them from financial stress.
- “Strivers”—strong psycho-social skills and sense of career progress but low income (personal and household) and dependent on public support.
- “Down and out”—very low psycho-social skills, low sense of career progress, severe life challenges, low income (personal and household), and strong reliance on public support.
- “Winners”—strong psycho-social skills and sense of career progress, high income (personal and household), few bill problems, and little dependence on either family or public support.

Missing College Credits

For missing credits, we assembled a rich set of predictors from the baseline forms (the PACE Basic Information Form/BIF and the Self-Administered Questionnaire/SAQ), NSC, the 18-month follow-up survey, person-level scales in the three-year survey, and spell-level data from the School Experiences module of the three-year survey. This was a spell-level file pooling data across the nine PACE sites, but not HPOG-only as no NSC data were available for the HPOG-only sample. We also added a large number of two- and three-way interactions with site and treatment group. After creating dummy variables for categorical variables, the total number of potential predictors was 1,584. The LASSO procedure working on this predictor set selected just six variables, yielding a model with an *R*-squared of 27 percent. Four of the six variables were significant predictors with standardized regression coefficients of at least 0.01. They were:

- Adjusted spell duration (adjusted for the longest break);
- Spell duration interacted with full/part-time student status;
- Credits reported at 18 months; and
- NSC-reported full-time-equivalent months of enrollment through 36 months after randomization.

After controlling on these factors, program and treatment were not important and nor were any of their interactions with each other or with other predictors. After imputing credits at the spell level, we summed to the person level for respondents with multiple school spells.

Missing Credential Award Dates

On the pooled PACE/HPOG credential sample, we modeled the lag between randomization and credential award date for those respondents with reported award dates ($n=12,392$, with 11,628 responses). The potential predictor list included site, treatment, the interaction of site with

treatment, type of credential (10 categories), life trajectory cluster, 20 parallel outcomes at the person level, the lag between randomization and interview, 16 baseline variables, and a large set of two- and three-way interactions with site and treatment group. After creating dummy variables for categorical variables, the total number of potential predictors was 1,160. The LASSO procedure working on this predictor set selected 14 variables, yielding a model with an *R*-squared of 8.4 percent. The significant predictors with standardized regression coefficients of at least 0.01 were:

- HPOG versus PACE;
- Credential was awarded for regular college classes and typically takes less than a year to earn;
- Credential is an associate degree;
- Credential is a bachelor's degree;
- Self-assessed career progress;
- Student debt;
- Three interactions of HPOG with main effects;
- One interaction of treatment status with a main effect; and
- Three 3-way interactions of HPOG status with treatment status with main effects.

After matching nonrespondents with respondents, we adjusted for the difference in randomization dates between the two people, by adding the lag from the respondent to the randomization date for the nonrespondent. If this was past the interview date for the nonrespondent, we truncated the award date to equal the interview date.

Missing Income

The instrument yielded four related measures of income in the past month: (1) exact personal income; (2) categorical personal income; (3) exact household income; and (4) categorical household income. As could be seen in Exhibit B-6, missing data rates were considerably higher for the continuous variables than the categorical variables. This is because categorical income is only missing if both exact (which can be put in the appropriate income category) and categorical income are missing. For prediction purposes, we assembled a person-level file with program, treatment status, the interaction of program with treatment status, self-reported earnings by quarter, 10 variables about economic well-being, four variables about psycho-social skills, nine measures of educational progress, 12 baseline characteristics, and a large collection of two- and three-way interactions with site and treatment group. We used this list for modeling both personal and household income. We ran the LASSO on the pooled PACE/HPOG three-year dataset ($n=14,467$, with 12,782 exact personal income reports and 9,219 exact household income reports). After creating dummy variables for categorical variables, the total number of potential predictors was 1,414.

The LASSO procedure working on this predictor set selected 10 variables for personal income, yielding a model with an *R*-squared of 58 percent. The significant predictors with standardized regression coefficients of at least 0.01 were:

- Dummy variables for three of the five life trajectory clusters;
- Personal earnings for the 12th quarter after random assignment;
- A dummy variable for having earned an associate degree since randomization;
- A scale for being able to make ends meet at the end of the month; and
- An interaction of earnings with a dummy for receipt of any means-tested public benefits.

For household income, the LASSO procedure selected 24 variables, yielding a model with an *R*-squared of 52 percent. The significant predictors with standardized regression coefficients of at least 0.01 were:

- Dummy variables for three of the five life trajectory clusters;
- Personal earnings for the 12th quarter after random assignment;
- A dummy variable for being an Earned Income Tax Credit claimant;
- A dummy variable for living with a spouse;
- A dummy variable for living with parents;
- A dummy variable for living alone;
- Annual baseline family income below \$15,000;
- Baseline SNAP (Supplemental Nutrition Assistance Program) or WIC (Special Supplemental Nutrition Program for Women, Infants, and Children) receipt;
- A dummy variable for having earned an associate degree since randomization;
- A scale for being able to make ends meet at the end of the month;
- An interaction of earnings with a dummy for receipt of any means-tested public benefits;
- An interaction of personal earnings with living arrangements; and
- Three 2- and 3-way interactions involving program.

Note that neither the model for personal income nor the model for household income involves three-way interactions of program with treatment status that are both statistically significant and substantively large. This does not mean that there are no program effects on income. Rather, it means that the measured parallel outcomes already capture whatever program effects might be present.

Certifications and Licenses in the First 18 Months

As mentioned earlier and as is discussed in detail in Section B.4 below, measures of ever-receipt of certifications and licenses blended reports from the 18 and three-year surveys. This decision also required imputing what nonrespondents²⁷ to the 18-month survey would have

²⁷ Nonrespondents here were people who could not be located, refused to be interviewed, or were otherwise unavailable for an interview. The concept does not include people who skipped questions about credentials when interviewed at 18 months. We assumed that these respondents did not earn any credentials by the time of the 18-month interview.

reported if they had responded at that time. We used the core imputation described above for this imputation.

On the pooled PACE three-year survey respondent sample ($n=6,773$ people, of whom 5,910 responded to both the 18-month and three-year follow-up surveys and 863 responded to only the three-year survey), we modeled the receipt of such credentials among those who responded to the 18-month follow up. The potential predictor list included program, treatment status, the interaction of program with treatment status, and about 40 baseline and three-year follow-up variables. After creating dummy variables for levels of categorical variables, this led to 80 potential predictors in total.

The LASSO selected 12 of the 80 predictors, including treatment status, dummy variables for three programs, one treatment-by-program interaction, six measures of educational progress and well-being at three years, and a dummy variable for employment in healthcare at three years. Of these, the predictors with large standardized coefficients included the count of exam-based certifications and licenses as of the time of the three-year survey, report of a short-term college credential at three years, current employment in healthcare, and the interaction of treatment with the Bridge to Employment program. This model achieved an R -squared of 12.0 percent, a high value for a binary outcome.

After imputing new exam-based certifications and licenses for 18-month survey nonrespondents, we separated exam-based certifications and licenses reported in the three-year survey using the donor's interview date into two categories—early (would have been reported by the nonrespondent in the 18-month survey if the interview had taken place) versus late (would have been earned after the 18-month survey if the interview had taken place). We then created a blended flag for having earned an exam-based certification or license as of the three-year survey. The flag was set to yes if the 18-month nonrespondent had an imputed early exam-based certification or license or had reported a late exam-based certification or license in the three-year survey.

B.2.2 Skipout, Start and End Dates, and Unreported School Spells

The remaining three types of missing data required more customized procedures. This section provides details on the approach to each type.

Skipout

We considered several approaches to this type of missing data. One option we considered and rejected was to treat respondents with skipouts as nonrespondents and give them nonresponse-adjusted weights of zero. However, this would have significantly boosted the overall nonresponse rate and wasted information collected after the skipout. A second rejected approach would have been to treat respondents with skipouts as nonrespondents only for analyses involving educational progress and employment. This option would have required the creation of a second set of nonresponse-adjusted weights, and would have led to inconsistencies across analyses. A third rejected option was to impute each outcome and scale requiring any data from the Integrated Training and Employment History module. This option was more attractive but would not have supported estimation of career trajectories.

The approach we adopted was to use a block imputation approach that was initially used in medical expenditure surveys in the United States (Williams and Folsom 1981). The general method involves matching a nonrespondent to a respondent and then copying the entire block of missing data from the respondent to the nonrespondent. Our objective was to find a respondent whose training and employment history would align well with the nonrespondent's baseline characteristics and measures of well-being at three years. (If the matched person had a missing response to a question within the Training and Employment History module, we copied this missing value over the skipout along with all the other variables.)

We used sequential hotdecks as in the core imputation methodology, but we formed the partitions in a different manner. Rather than modeling a single variable and then forming a nested set of partitions based on model-based predictions of that single variable, we crossed the life trajectory clusters discussed above with other important measures. We used a sequence of four hotdecks, where the first had the most stringent criteria for matches, and each succeeding hotdeck had loosened criteria.

The first hotdeck matched nonrespondents to respondents within cells defined by program, treatment status, any schooling reported prior to skipout, any work reported prior to skipout, life trajectory cluster, and lag between randomization and interview in whole months. This was on the pooled PACE/HPOG sample ($n=14,169$, with 13,245 respondents who did not skip out).²⁸ This run found donors for 815 of the 924 skipouts on the pooled dataset.

The second hotdeck replaced program with site. This run found donors for 86 of the remaining 109 skipouts on the pooled dataset. The third hotdeck replaced the exact number of months in the reference period with a dummy variable for whether the number was greater than 38 months. This run found donors for 22 of the remaining 23 skipouts on the pooled dataset. The fourth hotdeck used a collapsed version of self-assessed goal progress in place of life trajectory cluster and the binary recode of length of the reference period. This found a donor for the last remaining skipout.

Given the challenges in matching many of the nonrespondents to appropriate respondents, we did not carry out multiple imputation for skipouts. For the imputation of skipouts, our judgment was that the donor pools would be frequently small and that multiple random matches would, in fact, be the same match over and over. This lack of variation in the matched donors would have rendered variance estimates based on multiple imputations little better than variance estimates based on single imputation.

Because respondents with skipouts were missing a very long stretch of data that are important to most of the secondary outcomes in this report, we prepared impact estimates with and without these cases, as displayed in Exhibit B-7 below. The two sets of impact estimates are very similar. The imputation allowed us to use as many as 56 more cases for Bridge to Employment, about a 9 percent increase, with the exact count depending on item nonresponse.

²⁸ This excludes 302 three-year survey respondents that reported no training or employment between randomization and the survey interview.

Imputation shifted the impact of the program most on full-time-equivalent months of enrollment in any school, but not enough to change the statistical significance of the impact. That is, with or without this imputation, there is no evidence for a program impact on that outcome.

Exhibit B-7: Comparison of Selected Impact Estimates of Bridge to Employment with and without Skipouts

Outcome and Sample	Impact Estimate	Standard Error	Sample Size	p-Value
Employed at Survey Follow-Up (%)				
Full sample	4.2	3.7	658	.127
Omitting skipouts	2.7	3.9	602	.241
Employed at \$15 Per Hour or Above (%)				
Full sample	-1.4	3.6	643	.648
Omitting skipouts	-1.5	3.8	588	.650
Employed in a Job Requiring at Least Mid-Level Skills (%)				
Full sample	-3.7	3.5	641	.856
Omitting skipouts	-3.4	3.6	588	.826
Employed in the Healthcare Field				
Full sample	9.6***	4.0	654	.009
Omitting skipouts	10.5***	4.2	598	.006
Full-Time-Equivalent Months Enrolled in Any Type of School (months)				
Full sample	0.10	0.72	655	.444
Omitting skipouts	0.24	0.75	599	.374
Receipt of a Healthcare Credential from Any Type of School (%)				
Full sample	15.5***	3.9	658	< .001
Omitting skipouts	16.0***	4.1	602	< .001
Receipt of an Exam-Based Credential (%) (Unblended Version)				
Full sample	8.3***	2.9	658	.002
Omitting skipouts	8.2***	3.1	602	.004

Source: National Directory of New Hires; National Student Clearinghouse, and the PACE three-year follow-up survey.

Note: "Full sample" rows include values imputed for skipouts. All estimates are regression-adjusted as discussed in Appendix Section A.3. Statistical significance levels, based on one-tailed *t*-tests tests of positive differences between research groups for positive outcomes and negative differences for negative outcomes (such as student debt), are summarized as follows: *** 1 percent level; ** 5 percent level; * 10 percent level.

Spell Start and End Dates

As mentioned earlier, respondents were frequently unable to remember dates. We decided to impute them to make the most use of the partial information in each respondent's reported history. Our primary objective was to create a high-quality measure of the duration of study over the entire reference period. Secondary objectives included the ability to estimate quarterly earnings over the entire reference period and supporting a broader set of exploratory analyses of career trajectories (transitions between school, work, and other activities).

For this imputation, we used a different approach from any of those discussed above. This decision was motivated by the complexity of partial information in the Training and Employment

History module. Across the pooled PACE/HPOG sample, respondents had as many as six school spells and as many as 11 job spells. Even when respondents could not remember dates, we had many bounding conditions (e.g., spell 4 started after spell 3 ended). We devised a method that would respect these bounding conditions so as to create a coherent history while also supporting high-quality estimates of the site-specific impact of treatment on duration of study and quarterly earnings.

Before explaining the method, it will be useful to have an understanding of bounding conditions.

- For every spell, we knew whether it ended before the three-year follow-up interview or was ongoing at that time.
- For all closed spells, we knew whether there was another spell that started after it but prior to the three-year interview.
- For most spells, we knew
 - whether it started before or after randomization;
 - whether it started in the middle of another spell or after some period during which the person was neither working for pay nor enrolled in school; and
 - whether a new spell started during it.
- For spells that followed other spells, we would most often know the end date of the prior spell.
- For spells that preceded other spells, we would most often know the start date of the succeeding spell.
- For spells that started during other spells, we would most often know the start and end dates of the “mother” spell.
- For spells that spanned the start of a new spell, we would most often know the start and end dates of the “daughter” spell.

Our general approach to imputing missing dates involved the following steps on the pooled PACE/HPOG sample:

1. Express the date as a lag to some benchmark date. Specifically, we expressed start dates of main spells (those that did not start in the middle of any other spell) as the lag between randomization and the start of the spell, start dates of daughter spells as the lag from the start of the mother spell to the start of the daughter spell, and end dates of all spells as the lag from spell start date to spell end date.
2. Construct a statistical model for lag, and extract the predicted lag for spells with both known and unknown dates. (More details on this modeling process follow below. We constructed nine separate models.)
3. Identify the nearest neighbor case in the pooled dataset in terms of the predicted lag. Copy the lag from the spell with the known relevant date (start or end) to the case with an unknown value for the relevant date.

4. Add the imputed lag onto the benchmark date for the spell with an unknown date to obtain a preliminary date.
5. If the preliminary imputation violates any of the constraints, truncate it to just barely satisfy the constraints. For example, if preliminary imputation of an end date placed the end date past the date of follow-up interview but the respondent had reported that the spell ended before the interview, then we truncated the lag so that the job ended the month before the interview.

Before providing details on the nine models constructed in step 2, we offer some general observations about this methodology. We gave consideration to conducting this process separately for each site. We rejected that approach because of the complexity of the boundary constraints on dates and the rarity of patterns for respondents with multiple spells. Instead, we focused on constructing high-quality models and then finding the best match available.

The pooled sample size consisted of 27,939 job spells plus 13,093 school spells. After discarding spells reported by skipouts and spells that ended prior to randomization, the total number of spells was 40,672. Among these spells, either the start date or the end date was missing for 3,302, or 8 percent. Missing start dates was the more common problem, with 538 spells missing just the end date and 2,764 missing just the start date or both dates. Missing dates were slightly more common for school spells than for job spells (10 percent versus 7 percent). Missing dates for closed spells were much more common than for open spells (10 percent versus 4 percent). For Bridge to Employment, the overall missing data rate was slightly higher than for the rest of the pooled sample (11 percent versus 8 percent).

Exhibit B-8 below lists the models we created for each type of lag and some features of each, including average imputed values for the various lags. Main spell #1 was always the ongoing spell at the time of randomization for those respondents working or going to school at the point of randomization, and so always has a negative lag. Main spell #2 was always the first spell after randomization for those not working or going to school at the point of randomization. Other main spells always followed main spell #1 or #2. Given this structure, we prepared separate models for the start date of each group (lag types 1, 5, and 6 above) and we modeled other features associated with the first spell separately, as well (lag types 2, 3, and 4).

The set of variables allowed into each model varied across the nine lag types. Tested variables included program, randomized treatment group, the interaction of program with treatment group, elapsed time between randomization and follow-up interview (and its square), job/school status, next activity (work, school, or other), school control (three levels, nested within job/school status), school level (three levels, nested within job/school status), open/closed status, life trajectory cluster (five levels), self-assessed goal progress, baseline covariates, two- and three-way interactions of these variables with program and treatment status, and other variables.

Model fit as measured by *R*-squared varied substantially across models, ranging from 0 percent to 79 percent. The reasons for this variation are not clear to us. Average imputed values were generally quite similar to average reported months.

Exhibit B-8: Date Imputation for Three-Year Impact Study (Pooled Sample)

Lag Type	Modeled Variable	R-Squared (%)	Tested Variables	Selected Variables	Sample Size	Missing Data Rate (%)	Average Lag/Duration	
							Reported (months)	Imputed (months)
1	Lag from randomization date to start of main spell #1 (always negative because spell #1 was activity at time of randomization)	15	1,071	18	8,994	9.7	-18.8	-18.6
2	Duration of main spell #1 (closed only)	79	3,625	3	7,377	7.3	26.7	28.0
3	Lag from start of main spell #1 to start of subspell	78	2,989	3	5,459	8.8	23.2	16.9
4	Duration of subsPELLs of main spell #1 (closed only)	0	3,103	2	4,563	8.8	16.2	15.7
5	Lag from randomization date to start of main spell #2	7	1,089	2	3,610	7.0	6.7	6.7
6	Lag from randomization date to start of main spells #3 and higher	38	5,113	33	18,082	4.9	18.9	17.4
7	Duration of main spells #2 and higher (closed only)	16	4,760	23	13,509	5.4	8.3	8.3
8	Lag from start of main spell #2 and higher to start of subspell	43	4,105	11	4,270	6.3	6.0	4.2
	Duration of subsPELLs for main spells #2 and higher (closed only)	14	3,383	9	2,546	6.8	7.3	7.1

Source: National Directory of New Hires; National Student Clearinghouse; PACE three-year follow-up survey.

Note: Sample pooled across HPOG 1.0 and all nine PACE sites. ²⁹ Sample also pooled across treatment and control samples. A "main spell" is a spell that did not start in the middle of another spell. A "subspell" is a spell that did start in the middle of another spell.

Undercoverage of NSC-Reported Spells

As noted previously, we decided to supplement the histories of survey respondents who reported no training since randomization with any spells recorded for them in the NSC and then to impute the spell attributes collected in the survey beyond the simple start and end dates for the spells. Across the nine PACE sites, this edit changed the training history for 7 percent of the sample, switching them from a status of no training to some. In the Bridge to Employment sample, there were 33 such respondents, accounting for 5 percent of the sample. We added these NSC-reported spells to the three-year follow-up survey history for those respondents and imputed the missing survey outcomes such as earned credits and credentials.

This imputation proceeded by matching these 33 respondents to other Bridge to Employment study participants and copying over the donors' outcomes. This matching was structured, not

²⁹ ACF's Health Profession Opportunity Grants (HPOG) Program, like PACE, provides training to low-income individuals, but only for healthcare occupations. The impact study of 32 first-round HPOG awardees (HPOG 1.0) included three awardees also studied in PACE. For more: <https://www.acf.hhs.gov/ofa/programs/hpog>.

random. We constrained matches to be from the same treatment group and to have a similar predicted profile of four survey-reported spell-level variables:

- Received a diploma or certificate typically requiring less than a full year's worth of study during the spell;
- Received a diploma or certificate typically requiring a year or more's worth of study, but less than an associate degree during the spell;
- Received an associate degree or higher during the spell; and
- Total credits earned during the spell.

We formed linear models for each of these survey-reported spell-level outcomes in terms of baseline variables and NSC-reported spell- and person-level variables on enrollment and credential attainment. We fit these models on the pooled (treatment plus control) sample for the Bridge to Employment program. Given that the matching was not random, we did not conduct multiple imputation. We instead conducted single imputation and have ignored the impact on variances.

Of the 33 respondents with an added school spell, the imputation procedure resulted in 15 of them having a credential that they themselves did not report. Exhibit B-9 compares estimated program impacts with and without the addition of NSC-reported spells for respondents with no reported school spells since randomization. The two sets of impact estimates are very similar.

Exhibit B-9: Comparison of Selected Impact Estimates of Bridge to Employment with and without Imputation of NSC-Inferred Unreported Spells

Outcome and Sample	Impact Estimate	Standard Error	Sample Size	p-Value
Full-Time-Equivalent Months Enrolled in Any Type of School (months)				
Full sample	0.1	.7	655	.444
Omitting NSC-only spells	0.0	.7	655	.478
Receipt of a Healthcare Credential from Any Type of School (%)				
Full sample	15.5***	3.9	658	< .001
Omitting NSC-only spells	16.1***	4.0	619	< .001

Source: National Directory of New Hires; National Student Clearinghouse; PACE three-year follow-up survey.

Note: All estimates are regression-adjusted as discussed in Appendix Section A.3.

Statistical significance levels, based on one-tailed *t*-tests tests of positive differences between research groups for positive outcomes and negative differences for negative outcomes (such as student debt), are summarized as follows: *** 1 percent level; ** 5 percent level; * 10 percent level.

B.3 Survey Nonresponse Analysis

As in any survey, nonresponse can lead to bias if nonresponse propensity is correlated with outcomes. In the context of a randomized experiment such as this evaluation of Bridge to Employment, concern about nonresponse is heightened if the nonresponse rate is different in the treatment group than in the control group. Nonresponse can lead to biased impact estimates

even without differential nonresponse rates across study groups, but it is widely accepted that differential rates heighten concerns about biased impact estimates.³⁰

The three-year follow-up survey for this PACE site obtained a somewhat higher response rate in the treatment group (68 percent) than in the control group (63 percent) for the Bridge to Employment sample. We studied this matter further using administrative data and concluded that the nonresponse was causing serious bias in estimated impacts. (Illustrations of these biases are presented in Exhibit B-11 below.) We developed a set of nonresponse adjustment weights that appears to remove most of this bias. This section first presents the evidence of nonresponse bias in unadjusted impact estimates and then documents the nonresponse adjustment weights that we created to mitigate this bias.

B.3.1 Evidence of Nonresponse Bias in Unadjusted Impact Estimates

We gauged the likelihood of nonresponse bias through two types of analysis, one involving baseline data and one involving post-randomization administrative data.

The first analysis takes baseline equivalence as an indication of the potential for bias. If randomization is correctly implemented, there should be no systematic differences between the treatment group and the control group. We directly tested that using complete data from the BIF (see Appendix Section A.2). This insight also provides a proxy for nonresponse bias and the ability of our weighting scheme to correct for it. In the absence of nonresponse bias, appropriately weighted tabulations of the BIF among survey respondents should also show baseline equivalence.

The second type of analysis looks directly at estimated impacts. We know who responded to the survey and we have administrative data outcomes for both survey respondents and nonrespondents. We can thus compute two impact estimates from the administrative data: one estimate from the unweighted full sample, which we treat as truth; and a second estimate from the weighted survey sample. In the absence of nonresponse bias (and with large enough samples), we should get the same (up to sampling variability) estimates of impact on the full sample and on the weighted sample of survey respondents. Theoretically, it is possible to test whether estimated differences between these two impact estimates are statistically significant, but we did not do this, relying instead on impressions of consistency across a collection of administratively measured outcomes.

Exhibit B-10 below considers baseline equivalence, among survey respondents, in the absence of weighting.³¹ There were five significant imbalances (highlighted in red, using a threshold for statistical significance of 10 percent) on the full sample and seven significant imbalances on the unweighted survey respondent sample. By itself, this would not be much cause for concern. Comparisons of impacts on administrative follow-up outcomes with and without survey (as shown in Exhibit B-11) show much stronger evidence of nonresponse bias.

³⁰ See for example, Deke and Chiang (2017). For a slightly contrarian view, see Hendra and Hill (2018).

³¹ Note that the numbers in the first three columns of Exhibit B-10 reflect baseline balance for the full sample following imputation, whereas Appendix A.2 presented pre-imputation figures.

Before discussing Exhibit B-11, we note that the final column of Exhibit B-10 shows how baseline balance is improved by the weights. After application of the weights, only four baseline characteristics are out of balance, better even than on the full sample.

Exhibit B-10: Baseline Balance on Full Sample, Unweighted Respondent Sample, and Weighted Respondent Sample

Characteristics	All Participants			Survey Respondents, Unweighted			Survey Respondents, Weighted		
	Treatment	Control	p-Value	Treatment	Control	p-Value	Treatment	Control	p-Value
Age (%)			.252			.791			.660
20 or under	10.5	14.3		10.4	11.2		11.0	13.9	
21-24	19.4	20.7		17.9	20.5		17.3	18.7	
25-34	33.6	31.1		32.7	30.1		32.8	30.3	
35+	36.6	33.9		39.0	38.1		38.9	37.1	
Gender (%)			.689			.489			.922
Female	83.2	84.1		83.0	84.9		82.5	82.8	
Male	16.8	15.9		17.1	15.1		17.5	17.2	
Race/Ethnicity (%)			.760			.285			.366
Hispanic, any race	48.0	47.79		52.60	49.36		48.1	51.3	
Black, non-Hispanic	21.5	20.88		18.21	20.19		19.8	19.6	
White, non-Hispanic	18.6	21.08		16.18	22.44		17.7	20.9	
Other, non-Hispanic	15.0	14.66		15.32	13.46		16.8	12.9	
Family Structure (%)			.797			.737			.566
Not living with spouse/partner and not living with children	46.3	46.6		43.01	45.8		44.3	47.5	
Not living with spouse/partner but living with children	29.1	26.9		31.8	27.9		31.0	26.2	
Living with spouse/partner and not living with children	11.9	11.9		12.1	12.2		12.2	11.9	
Living with spouse/partner and children	12.9	14.7		13.0	14.1		12.5	14.4	
Living with parents (%)	26.5	31.7	.067	28.0	30.1	.555	28.9	31.2	.517
One parent has at least some college (%)	46.4	42.0	.154	43.4	42.3	.787	44.1	41.2	.450
High School Grades (%)			.615			.461			.882
Mostly A's	21.3	18.9		22.3	19.2		20.6	19.4	
Mostly B's	53.2	55.2		48.0	52.6		53.3	53.0	
Mostly C's or below	25.5	25.9		29.8	28.2		26.1	27.7	

	All Participants			Survey Respondents, Unweighted			Survey Respondents, Weighted		
Current Education (%)			.261			.035			.038
Less than high school diploma	4.7	2.6		5.8	1.9		5.5	2.0	
High school diploma or equivalent	34.6	39.2		33.0	37.8		31.9	40.5	
Less than 1 year of college	19.6	18.7		18.5	20.8		18.7	19.0	
1 or more years of college	24.5	22.1		25.1	19.6		26.0	21.2	
Associate degree or higher	16.6	16.9		17.6	19.6		17.8	17.2	
Received vocational or technical certificate or diploma (%)	45.3	43.6	.592	43.9	45.8	.625	44.1	45.4	.734
Career Knowledge Index (average of items)	0.59	0.52	.004	0.58	0.52	.044	0.58	0.51	.010
Psycho-Social Indices (average of items)									
Academic Discipline Index	5.59	5.52	.032	5.58	5.5	.037	5.59	5.49	.021
Training Commitment Index	5.74	5.71	.242	5.74	5.73	.66	5.74	5.72	.600
Academic Self-Confidence Index	5.14	5.08	.166	5.12	5.07	.368	5.11	5.06	.304
Emotional Stability Index	5.41	5.38	.342	5.41	5.38	.456	5.41	5.37	.424
Social Support Index	3.37	3.34	.146	3.38	3.33	.071	3.38	3.33	.102
Stress Index	2.11	2.17	.137	2.10	2.18	.114	2.1	2.17	.217
Depression Index	1.43	1.47	.098	1.42	1.48	.064	1.42	1.47	.111
Income (%)			.177			.195			.259
Less than \$15,000	50.4	57.2		48.3	55.8		47.9	54.6	
\$15,000-29,999	30.2	26.9		31.5	27.9		32.3	27.9	
\$30,000+	19.8	16.9		20.5	17.3		20.2	18.6	
Mean (\$)	17,750	17,575	.866	18,314	17,603	.572	18,338	17,957	.764
Public Assistance / Hardship Past 12 Months (%)									
Received WIC or SNAP	46.4	49.2	.383	48.6	52.6	.305	48.1	48.9	.854
Received public assistance or welfare	20.6	18.7	.454	19.1	19.2	.960	19.1	18.4	.825
Reported financial hardship	51.2	56.4	.096	52.0	60.9	.022	52.1	57.5	.167

	All Participants			Survey Respondents, Unweighted			Survey Respondents, Weighted		
Current Work Hours (%)			.806			.813			.896
0	63.2	60.6		63.6	62.2		62.8	61.1	
1-19	9.7	10.6		10.4	10.9		9.8	10.4	
20-34	15.4	17.3		13.9	16.0		14.6	16.5	
35+	11.7	11.2		12.1	10.6		12.7	11.6	
Expected Work Hours in Next Few Months (%)			.748			.558			.543
0	22.7	25.5		23.1	26.3		23.4	26.2	
1-19	9.7	8.8		11.3	8.7		11.3	8.7	
20-34	29.8	29.7		29.8	27.9		30.0	27.4	
35+	37.8	35.9		35.8	37.2		35.3	37.7	
Life Challenges Index (average in original units 1-5)	1.42	1.49	.010	1.42	1.52	.009	1.42	1.50	.020
Owns a car (%)	69.4	67.1	.434	69.1	68.9	.964	69.1	69.2	.958
Has both computer and internet at home (%)	74.1	75.9	.512	73.4	76.9	.299	73.9	76.7	.405
Ever arrested (%)	13.4	13.9	.848	13.0	14.4	.598	12.7	14.5	.499
Sample sizes	506	498		346	312		346	312	

Source: PACE Basic Information Form; PACE Self-Administered Questionnaire; response status to the PACE three-year follow-up survey. Sample restricted to those randomized for the evaluation of Bridge to Employment.

Note: SAS/SURVEYFREQ used to test for significant imbalances for categorical variables. SAS/TTEST used to test for significant imbalances for other variables. Weights are based on the dual raking system explained in Appendix Section B.3.2 below. Significant imbalances are highlighted in red, using a threshold for statistical significance of 10 percent. Statistical significance was measured using two-tailed tests for differences.

Exhibit B-11 presents evidence about the level of nonresponse bias with and without adjustment weights. The first two panels compare three sets of regression-adjusted impacts on earnings outcomes from NDNH records (panel 1) and on college outcomes from NSC records (panel 2).³² The first set of impact estimates (column 1) is based on the full sample. The second set (column 3) excludes survey nonrespondents. Differences between the first and second set of impacts signal nonresponse bias. The third set of impacts (column 5) also excludes survey nonrespondents but weights survey respondents with nonresponse adjustment weights, which are explained in Section B.3.2 below. If the weights are good, then the differences between the first and fifth columns will be smaller than those between the first and third columns. Note that all three impact estimates are regression-adjusted with the covariates discussed in Appendix Section A.3.

Exhibit B-11: Comparison of Selected Estimates of the Impact of Bridge to Employment for the Unweighted and Weighted Survey Samples

Outcome (Data Source)	Full Sample		Unweighted Sample		Weighted Sample	
	Impact Estimate	Standard Error	Impact Estimate	Standard Error	Impact Estimate	Standard Error
Confirmatory Outcome (NDNH)						
Quarterly earnings (average of 12th and 13th quarters after randomization) (\$)	289	254	738**	319	340	331
Q5 earnings (\$)	8	210	201	262	-55	274
Q9 earnings (\$)	184	241	486*	296	110	318
Q13 earnings (\$)	200	272	615**	340	254	359
Q17 earnings (\$)	106	319	418	399	105	415
Any earnings Q5 (%)	-1.7	2.9	-0.9	3.7	-3.4	3.7
Any earnings Q9 (%)	1.8	2.8	3.0	3.5	-0.1	3.5
Any earnings Q13 (%)	0.4	2.8	1.5	3.4	-0.8	3.4
Any earnings Q17 (%)	-0.5	2.7	0.3	3.4	-1.6	3.4
Auxiliary Education Outcomes (NSC)						
Number of months of enrollment through 35 months	-1.2	0.6	-1.9	0.7	-1.2	0.7
Number of months of full-time enrollment through 35 months	-0.7	0.4	-1.3	0.5	-0.7	0.5
Any enrollment through 35 months (%)	-5.6	3.0	-10.4	3.7	-6.0	3.8
Any credentials through 35 months (%)	-2.2	1.4	-1.8	1.7	-2.5	1.8
Number of months of full-time enrollment through Oct 2018	-0.9	0.7	-1.6	0.9	-0.6	0.8
Any credentials through Oct 2018 (%)	-1.6	1.9	-2.4	2.3	-2.0	2.3

³² The NSC outcomes in this table are not formal outcomes for the evaluation of Bridge to Employment. We decided not to use them for the formal evaluation because many students use their vouchers at schools that do not report to the NSC. Nonetheless, these outcomes are observed for the full sample and thus are useful for assessing the contribution of the weights to inference.

Outcome (Data Source)	Full Sample		Unweighted Sample		Weighted Sample	
	Impact Estimate	Standard Error	Impact Estimate	Standard Error	Impact Estimate	Standard Error
Secondary Employment Outcomes (Survey)						
Employed at survey follow-up (%)			6.2**	3.6	4.2	3.7
Employed at \$15 per hour or above (%)			0.6	3.5	-1.4	3.6
Employed in a job requiring a least mid-level skills (%)			-3.0	3.4	-3.7	3.5
Employed in the healthcare field (%)			11.3***	3.9	9.6***	4.0
Secondary Education Outcomes (Survey)						
Full-time-equivalent months enrolled at any school			-0.7	0.7	0.1	0.7
Receipt of a healthcare credential from any school (%)			16.9***	3.8	15.5***	3.9
Receipt of an exam-based credential (%)			24.8***	3.8	22.0***	4.0
Other Secondary Outcomes (Survey)						
Indicators of Independence and Well-Being						
Health insurance coverage (%)			2.6	2.7	3.0	2.8
Receives public benefits (%)			-0.3	3.7	1.1	3.8
Student debt (\$)			-1,650	760	-919	695
Financial hardship (%)			0.9	3.8	2.1	3.8
Indices of Self-Assessed Career Progress (average)						
Confidence in career knowledge ^b			0.02	0.04	0.02	0.04
Access to career supports ^c			-0.02	0.02	-0.03	0.03
Sample sizes (across treatment and control groups)	1,004		658		658	

Source: NDNH, NSC, and the PACE three-year follow-up survey.

Note: All estimates are regression-adjusted as discussed in Appendix Section A.3. The full sample columns are blank for survey-measured outcomes as they are not available for the full sample.

Statistical significance levels, based on one-tailed *t*-tests tests of positive differences between research groups for positive outcomes and negative differences for negative outcomes (such as student debt), are summarized as follows: *** 1 percent level; ** 5 percent level; * 10 percent level.

While we did not formally test the differences between the alternative estimates, given that the survey respondents constitute a very large subset of all participants, many of the differences would be statistically significant. For several follow-up administrative variables, there appears to be substantively important bias in estimated impacts based on the unweighted respondent sample. Most troubling, the estimated impact on the average of quarterly earnings for the 12th and 13th quarters is large and statistically significant on the unweighted respondent sample (\$738), whereas for the full sample the estimated impact is modest and not statistically significant (\$289). Also troubling, the estimated impact on any enrollment at an NSC-cooperating school is a 10.4 percentage point drop on the (unweighted) respondent sample. This would be statistically significant at the 1 percent level if we had run a two-sided hypothesis test. The estimated drop of 5.6 percentage points on the full sample is more modest and only be significant at the 10 percent level if we had run a two-sided hypothesis test.

The positive bias in the estimated program impact on earnings and the negative bias in the estimated program impact on school enrollment and perseverance led us to conclude that both current earnings and educational progress are related to nonresponse propensity in different ways on the treatment and control groups. Given the centrality of educational progress and earnings in the logic models for how Bridge to Employment would affect a wide variety of life outcomes measured in the survey, these relationships clearly imply some survey nonresponse adjustment is required.

The final pair of columns shows that the nonresponse weights bring impact estimates based only on survey respondents back into good alignment with impact estimates on the full sample. For example, the impact on quarterly earnings in the 12th and 13th quarters for the full sample is \$289. The estimate impact for the weighted survey sample is \$340, which is much closer to the full sample estimate than the unweighted estimate. Neither of these impact estimates are statistically significant. In contrast, the unweighted impact for the survey sample is \$738, and it is statistically significant. This illustrates how the nonresponse weights removed much of the bias in the unweighted survey sample. The weighted impacts do not agree exactly with the full-sample impacts, but that would be an unreasonable goal for an adjustment procedure.

We implemented this solution across all nine PACE sites. Nonresponse bias was worse for Bridge to Employment than at any of the other sites; still, unweighted analyses were subject to serious bias at three of the other sites, and the procedure appears to do no harm even when not strictly required.

For the survey-based outcomes, the three bottom panels of Exhibit B-11 compare the unweighted and weighted impact estimates. There are some differences between the estimates—most notably, the impact on follow-up employment is 6 percentage points and significant in the unweighted sample, but is smaller (4 percentage points) and not significant in the weighted sample. This is consistent with the NDNH findings, which found that the unweighted survey sample tended to upwardly bias employment and earnings impacts. There is also a sizeable difference in the impact on average student debt—the impact of –\$1,650 in the unweighted sample is reduced to –\$919 in the weighted sample.

B.3.2 Construction of Nonresponse Adjustment Weights

Construction of weights to reduce the biases just discussed was more complex than anticipated. At first, we tried a standard propensity scoring approach,³³ as was used in the short-term report on Bridge to Healthcare (Farrell and Martinson 2017). However, that approach was not successful in removing the biases shown in the unweighted column of Exhibit B-11. Data storage arrangements posed a further challenge in developing a set of nonresponse adjustment weights. Contractual arrangements permitted the merging of survey data with either NDNH data

³³ In the standard approach, a logistic model for response status is fit in terms of universally available covariates (baseline and administrative), the model is used to generate a predicted response propensity for each person (respondent and nonrespondent), people are sorted on this prediction into strata, the empirical response rate is calculated for each stratum, and finally the inverse of this rate is applied to respondents as a nonresponse-adjustment weight.

or NSC data, but they did not permit the merging of NDNH and NSC data. In response to this challenge, we developed a new approach that we call dual-system raking.

“Raking” is the name for iterative procedures that create weights for a sample in such a manner that marginal tabulations of the sample agree exactly with pre-specified “control” totals in multiple dimensions. For example, raking can be used to create weights that will cause tabulations by gender, tabulations by race, and tabulations by age all to agree with pre-specified totals for gender, race, and age. In this example, gender, race, and age are dimensions.

In the context of nonresponse, if tabulations are prepared from the full sample and raking is used on the respondents, then weighted tabulations of the respondent sample will be in perfect agreement with parallel tabulations of the full sample. This exact multi-dimensional agreement is referred to as “hyperbalance.” In the context of an experiment, if this procedure is run separately for the treatment and control groups, then hyperbalance between respondents and nonrespondents means that the weighted balance between the treatment and control groups on the respondent sample should be just as good as on the full sample.

This hyperbalance by arm means that if we estimated treatment impact on just the respondent sample with these weights but without regression adjustment, the estimated program impact on each of these hyperbalanced variables would agree exactly with corresponding program impacts estimated on the full sample. The use of regression adjustment to estimate program impacts (rather than simple mean difference between arms) means that this agreement will not be exact, but agreement should still be very good for hyperbalanced variables. Theoretically, it should also improve agreement (between impact estimates based on the full sample and impact estimates based on just the respondent sample) for a variety of related parallel outcomes.

Key raking variables include both categorical variables (e.g., any NSC-reported enrollment) and interval-valued variables (e.g., number of months enrolled in college according to NSC records). Including these interval-valued variables seems particularly important because many educational outcomes are associated with the length of study.

The need to include continuous variables in the raking is challenging because traditional raking algorithms work only with categorical variables. In contrast, the generalized raking we propose and use here can handle a mix of categorical and continuous variables.³⁴ For categorical variables, the procedure guarantees perfect correspondence between the respondent sample and full sample by arm on the distribution of the sample across the categories of each variable; for continuous variables, the procedure induces perfect agreement on the marginal means of each of them.

The generalized raking procedure of Folsom and associates is available in the WTADJUST procedure of SUDAAN. A similar procedure that only works for categorical covariates is the SAS

³⁴ Generalized raking is most fully developed by Folsom and Singh (2000), who in turn draw on work originally proposed by Folsom (1991), Deville and Särndal (1992), and Folsom and Witt (1994). Dual raking is similar to the approach of Judkins et al. (2007) that involves the use of raking to construct weights in quasi-experimental designs.

raking macro of Izrael et al. (2000). It was necessary to use both of these software packages because the analyses had to be run on two servers, one that had SUDAAN installed (at Abt) and one that did not (at ACF). We refer to our system as dual-system raking because it permits raking both to NDNH information and to NSC information though the two types of data reside on two different systems.

The details of the dual-system raking procedure are as follows:

1. We used SUDAAN/WTADJUST to develop survey weights on the Abt server that induced hyperbalance by arm for the means of four NSC variables. Two of these NSC variables were counts on months: months with any enrollment and months of full-time-equivalent enrollment. Two of the NSC variables were binary flags: any enrollment and any completions (credentials). All four of these variables were constrained to enrollment and completions within 35 months of randomization.
2. We merged the weights from step 1 with baseline data and follow-up survey data on the Abt server. We then passed these merged data through to a secure ACF server, where third-party ACF contractors merged our data with NDNH earnings data, removing personal identifiers from the merged dataset. We had verified that this set of NSC-adjusted weights provides nearly unbiased impact estimates for survey-based education outcomes, but after merging the weights with NDNH data, we discovered that these NSC-adjusted weights did not remove bias in survey-based impact estimates for earnings outcomes.
3. To remedy this, we used the Izrael-Hoaglin-Battaglia macro on the ACF server to rake the weights from step 1 in such a manner as to attain hyperbalance by arm on three categorized versions of NDNH earnings. Specifically, we obtained hyperbalance for a six-level categorization of earnings at Q12 and Q13, a five-level categorization of earnings at Q9, and a five-level categorization of cumulative earnings from Q1 through Q12.³⁵ We verified that these weights removed most of the nonresponse bias on estimates of program impacts on NDNH earnings when estimated from nonrespondents instead of from the full sample. This sensitivity analysis included the continuous versions of the variables used in the raking, as well as continuous earnings at Q5 and Q17 and binary indicators for any employment at Q5, Q9, Q13, and Q17.
4. We used the weights from step 3 on the ACF server to estimate (by arm) the distributions of survey-reported earnings. Specifically, we split Q12 earnings at \$0, \$6,000, and \$9,000; Q9 earnings at \$0, \$6,000, and \$9,000; and average quarterly earnings for Q1 through Q12 at \$3,000 and \$6,000. (The breaks for survey-reported earnings needed to be coarser than the breaks for NDNH earnings because of the smaller sample sizes in the respondent survey sample.)

³⁵ This process is also referred to as “binning.” We used more bins for the confirmatory outcome than for the exploratory outcomes. Reducing the number of bins generally speeds convergence and reduces the frequency of extreme adjustments.

5. We again used the Izrael-Hoaglin-Battaglia macro on the ACF server to rake the weights from step 1, but for this step we used the control totals from step 4 rather than the NDNH totals used in step 3. We then verified that these weights removed most of the nonresponse bias on estimates of program impacts on NDNH earnings when estimated from nonrespondents instead of from the full sample. (Exhibit B-11 shows the results.) These weights did not perform as well as the weights from step 3 in reducing nonresponse bias on the respondent sample, but the deterioration (not shown) was not very large.
6. We exported the 11 estimated totals from step 4 for each arm from the ACF server to the Abt server. (The data use agreement permits the transfer of tabulations; only the export of microdata is prohibited.)
7. We again used the Izrael-Hoaglin-Battaglia macro to rake the weights from step 1 to the control totals from step 4, but this time we did the raking on the Abt server rather than on the ACF server. We then merged these with NSC data on the Abt server and verified that these weights removed most of the nonresponse bias on estimates of program impacts on NSC outcomes when estimated from nonrespondents instead of from the full sample. (Exhibit B-11 above shows the results.)

B.4 Quality and Completeness of Exam-Based Credentials Reported in the Survey

Estimates of exam-based certifications and licenses based on the three-year follow-up survey were much lower than those based on the short-term survey at 18 months after randomization (Farrell and Martinson 2017). This is logically impossible; over time people can only gain certifications and licenses.

A review of the survey's skip patterns and wording identified three features in the design of the three-year instrument that might have led to the reporting of few credentials of this type:

- First, the three-year instrument allowed only respondents with some formal schooling since randomization to report exam-based certifications and licenses. However, people who learn skills on the job or through independent online study (such as YouTube tutorials) can sit for the exams for many certifications and licenses.
- Second, the wording for the three-year instrument strongly emphasized that “school-issued certificates” were not the same thing as “exam-based certifications and licenses.” We had introduced this language to ease confusion about the difference between credentials issued by schools and credentials issued by other authorities. However, because some schools serve as proxy administrators of exams for credentials that are actually issued by other authorities, it is possible that this emphasis wording led some people to report exam-based credentials as school-based credentials or to not report them at all.
- The third feature is just the greater passage of time. Respondents may not have renewed exam-based certifications and licenses or they might have discovered that the

credentials are less useful than anticipated, either of which could have reduced respondents' inclination to report older exam-based credentials.

Given this review, we decided that the short-term follow-up survey reports of early exam-based credentials earned are probably more accurate than the reports from the three-year survey. Accordingly, we decided to combine reporting for the two time periods. The composite measure of receipt of any exam-based credential since randomization was set to yes if the respondent either reported it in the 18-month survey or reported receiving such a credential in the three-year survey at a time point after the date of the 18-month survey interview. For the 14 percent of the sample who did not respond at 18 months, we imputed a response. When receipt dates were not reported in the three-year survey, we also imputed them. Both of these imputations are discussed above in Section B.3.

B.5 Quality and Completeness of School-Issued Credentials Reported in the Survey

As discussed in Section B.4 above, an initial analysis showed that the estimated levels of receipt of new certifications and licenses—which this report also refers to as exam-based credentials—were lower than those published in the short-term impact report (Farrell and Martinson 2017), which is a logical impossibility. We have already discussed our solution to that problem.

In this section, we review why we believe these problems did not affect survey reports of school-issued credentials seriously enough to warrant our not carrying out similar adjustments for them. We then present sensitivity analyses that compare the impact of Bridge to Employment on credential receipt in the first 18 months using the 18-month follow-up survey and the three-year follow-up survey. We preface this analysis with a review of the instrumentation employed by the two surveys.

School-Issued Credentials, Three-Year Survey. The procedure we used to measure receipt of school-issued credentials in the three-year survey rested fundamentally on two interview questions:

- I2. Since [MONTH AND YEAR OF RANDOM ASSIGNMENT], have you received a diploma, certificate, or academic degree for completing any regular college classes?*
- I3. Since [MONTH AND YEAR OF RANDOM ASSIGNMENT], have you received any diplomas or certificates from a school for completing any vocational training?*

One reason to have two separate questions was to try to separate the credentials issued by regular departments of community colleges from those issued by workforce training institutions. The latter, though affiliated with community colleges, frequently have admissions criteria, curricula, instructors, and administrators different from the core college programs that are accredited to grant degrees. Another reason for attempting to make this distinction was the theory that more of the credits required to earn sub-degree credentials awarded by core college departments might count toward degrees and therefore be more valuable to students.

Throughout this report, for ease of reference, we refer to credentials named in response to question I2 as “academic” credentials and those named in response to question I3 as “vocational” credentials. Future research may explore whether the credentials reported in response to these two questions have differential value. At present, this is unclear. Some common school-issued credentials such as Licensed Practical Nurse are reported in response to both questions.

The three-year survey asked respondents who answered yes to either question I2 or I3 an additional set of follow-up questions on the number of such credentials, their names, award dates, and issuing school,³⁶ but the follow-up questions were slightly different for the two types of credentials. The survey asked respondents reporting academic credentials to describe the level of effort required to earn the credential, classified into one of the following categories:

1. *A diploma or certificate requiring less than a full year’s worth of credit.*
2. *A diploma or certificate requiring a full year or more’s worth of credit (but less than an associate degree).*
3. *An associate degree.*
4. *Bachelor’s degree or higher.*

The survey did not ask a parallel question of respondents reporting vocational credentials. Instead, we examined the reported credential names and imputed a level of effort based on general knowledge of postsecondary education.

School-Issued Credentials, 18-Month Survey. The 18-month survey started with a global question asking whether the respondent had “*taken any classes or been in an instructional program of any kind anywhere, even for a short time?*” Those who answered yes were asked to list schools attended, starting from the most recently attended and working backwards from there if respondents attended multiple schools. For each school, respondents were asked about the types of classes that they took at the school. Specifically, interviewers asked whether they had taken each of the following:

1. *Classes to learn English as a second language? Do not count any classes providing regular college credit or occupational training.*
2. *Other classes to improve your basic reading, writing, or math skills or prepare for a high school equivalency or college placement test? Again, do not count any classes providing regular college credit or occupational training.*
3. *Classes providing regular college credit?*
4. *Classes providing occupational training, but not for college credit?*

³⁶ Rather than making the interviewer type the respondent’s open ended-response, the survey instrument provided a list of attended schools (previously reported) from which to make a selection as the school of issue. Respondents were also allowed to name a school not previously named.

5. Classes in other skills, such as how to succeed at school, work, or other areas of life? (Please include any such classes, whether for college credit or not.)

If respondents answered yes to the third type, then they were asked about their goal related to the course taking, and whether they reached the goal. Possible goals included earning a diploma/certificate requiring less than a full year's worth of credit, a diploma/certificate requiring a year or more's worth of credit but less than an associate degree, an associate degree, or a bachelor's degree or higher.

If respondents answered yes to the fourth type, then they were asked whether their goal in this course taking was to attain a credential, and if so, whether they reached the goal. For those who reported a vocational diploma or certificate, they were asked a question to separate vocational credentials into three categories by nominal level of required effort—less than 10 weeks, 10 weeks to less than a year, or one year or more.

The Decision Not to Adjust. As discussed above, for credential receipt across the entire three-year period, we decided to interweave the early receipt of a certification or license with the later receipt of such a credential. For other types of credentials, we decided to use only the three-year survey. We based this decision on analyses of data for the PACE program at Pima Community College (PCC), Pathways to Healthcare. We chose this site for the research because we had PCC college records and because the evaluation's processing of those records was further along than was processing at other PACE sites for which we had negotiated access to college records.

Analysis of PCC records showed that the three-year survey was more accurate than the 18-month survey. We focused on respondents who reported a school-issued credential in only one of the surveys, and then checked to see whether the PCC records confirmed issuance of that survey-reported credential. Among respondents who reported such a credential at 18 months but not at three years, PCC records confirmed this claim for just 35 percent. In contrast, among respondents who reported such a credential at three years but not at 18 months, PCC records confirmed this claim for fully 81 percent. For some reason, the 18-month survey instrument seems to have generated many more unverifiable credential claims than the three-year survey. For this reason, we decided to rely on the three-year survey without adjustment for school-issued credentials in all PACE sites.

Sensitivity Analyses. Exhibit B-12 below shows the differences between the two surveys for the evaluation of Bridge to Employment. For this exhibit, we used credential award dates in the three-year survey to identify credentials awarded in the first 18 months after randomization. Ideally, with this restriction on award dates, the two surveys should produce similar estimates. We did not formally test for differences between the two versions of the impacts. However, the strong correlation induced by being mostly measured on the same set of people means that even modest differences are probably statistically significant.

The two surveys were largely consistent on the award of longer term academic credentials. Both failed to find an effect. The story was different for the other types of credentials. Only the three-

year survey detected an effect on short-term academic credentials through 18 months, and only the 18-month survey detected effects on vocational credentials and exam-based certifications and licenses through 18 months. These differences mean that the decision based on analyses at PCC strongly influenced inferences for Bridge to Employment.

Exhibit B-12: Comparison of Selected Impact Estimates on Credential Receipt in the First 18 Follow-up Months for Bridge to Employment

Outcome, Sample	Impact Estimate	Standard Error	Sample Size	p-Value
Received an Academic Credential That Typically Requires Less Than a Year of Credits (%)				
18-month survey	-1.4	1.4	566	.839
3-year survey	13.5	3.7	658	<.001
Received an Academic Credential That Typically Requires At Least a Year of Credits (%)				
18-month survey	-1.4	1.8	566	.794
3-year survey	0.7	1.8	658	.343
Received a Vocational Credential (%)				
18-month survey	28.2	3.5	566	<.001
3-year survey	13.6	3.1	658	<.001
Received Certification or License (%)				
18-month survey	28.6	4.1	566	<.001
3-year survey	8.0	2.5	658	.001

Source: PACE 18-month and three-year follow-up surveys.

Note: For the three-year survey, early credentials are those that are reported as occurring before the month of the prior interview if there was a prior interview. Otherwise, early credentials are those reported within 18 months after randomization. All estimates are regression-adjusted as discussed in Appendix Section A.3.

Statistical significance levels, based on one-tailed *t*-tests tests of positive differences between research groups for positive outcomes and negative differences for negative outcomes (such as student debt), are summarized as follows:

*** 1-percent level; ** 5-percent level; * 10-percent level.

Appendix C: NDNH's Unemployment Insurance Wage Data

Through the 1990s, many social program evaluations relied on administrative earnings data provided by state Unemployment Insurance (UI) agencies. State agencies maintained these data, and privacy concerns sometimes precluded sharing with outside researchers. UI records have become more accessible since 1996 with the advent of a centralized national database—the National Directory of New Hires (NDNH). Among NDNH's virtues is that, unlike state data, it captures earnings for study participants who move to another state during the follow-up period.

The federal Office of Child Support Enforcement (OCSE) administers the NDNH. It contains new hire, quarterly wage, and UI information submitted by State Directories of New Hires, employers, and state workforce agencies. OCSE also supplements the state reports with records about earnings from federal civilian and military jobs (which are otherwise not covered by state UI data). Given this supplementation, the most important uncaptured earnings are any unreported tips, self-employment, firms' employment of independent contractors, and informal employment.³⁷

C.1 Data Collection Process

The primary purposes of the NDNH are to assist state child support agencies to locate noncustodial parents, putative fathers, and custodial parents to establish paternity and child support obligations and to enforce and modify orders for child support, custody, and visitation. It is also used by state UI agencies and the federal Social Security Administration to identify overpayments of benefits. However, subject to federal law, regulation, guidance, and other requirements to protect data privacy and security,³⁸ OCSE may disclose certain information contained in the NDNH to requesting local, state, or federal agencies for research likely to contribute to achieving the purposes of part A or part D of title IV of the Social Security Act. Part A governs the federal Temporary Assistance for Needy Families (TANF) program. Part D governs the state/federal child support program. Such disclosures may not include the names, Social Security numbers (SSNs), or other personally identifying information. If the disclosure is approved, the agency and OCSE must work together on the operational issues surrounding the technical and procedural aspects of the disclosure, such as mitigating the risks of identifiability and establishing appropriate data retention and disposition schedules of data files.

The Office of Policy, Research, and Evaluation and OCSE negotiated a memorandum of understanding (MOU) allowing access to NDNH data for the PACE evaluation. Among other provisions, the MOU dictates what self-reported data from study subjects may be merged with

³⁷ According to the U.S. Bureau of Labor Statistics, about 10 percent of workers are self-employed: <https://www.bls.gov/spotlight/2016/self-employment-in-the-united-states/home.htm>

³⁸ The legal authority for this disclosure for research purposes is contained in subsection 453(j)(5) of the Social Security Act and subsection 5507 of the Patient Protection and Affordable Care Act.

NDNH data, the computing environment where these merges are conducted and the data analyzed, and procedures for review of tables prior to release.

The PACE research team transmits match request files to OCSE quarterly. These match request files contain the names and SSNs of PACE study participants. OCSE verifies with the Social Security Administration that the reported SSNs belong to the named persons. For those SSNs that pass this test, OCSE copies NDNH records for that quarter and the preceding seven quarters to a secure folder³⁹ (Ordinarily, these records would be destroyed after two years.) These copied records have been stripped of all personal identifiers and a pseudo-SSN is added.

States are required to submit earnings records to OSCE within four months, but there are stragglers and corrections. To be safe, PACE analyses limit NDNH-based measures to time periods that ended at least six months prior to the extract date.

Once we are ready to analyze the collected data, we submit a “passthrough” file to OCSE containing a variety of PACE-assigned variables (such as treatment status and program ID) and self-reported variables (such as the baseline information described in Appendix A). OCSE then strips the personal identifiers out of the passthrough file and replaces the actual SSNs with the same pseudo-SSNs previously assigned to the archived wage records. The study then uses these pseudo-SSNs to merge program and self-reported data with NDNH quarterly wage data on ACF’s secure server in order to estimate program impacts on earnings and employment.

C.2 Data and Measures

Random assignment for Bridge to Employment started in June 2012 and ended in October 2013. Given the lag of up to six months in processing of employer reports by the states and transfer of state data to OCSE, wage records from NDNH were available through Q2 2018; this means that we had 24 post-randomization quarters of earnings data for the earliest randomized study participants and 18 post-randomization quarters of earnings data for the last randomized study participants. In addition, we had eight quarters of pre-randomization data for the entire sample (we included only the four most recent of these quarters in our regression-adjustment models).

Of the 1,004 treatment and control group members randomized as part of the Bridge to Employment evaluation, 974 study participants reported a name and SSN that OCSE deemed to be of sufficient quality for its matching purposes.⁴⁰ Analyses in this three-year report thus are based on the 97 percent of the sample the agency deemed suitable. This sample’s earnings in each quarter were based on earnings records found for each sample member in matching. As usual in use of such data, we defined sample members as “not working” when there was no match to wage records in a given quarter.

³⁹ Those study participants who are not matched in the Social Security Administration database are considered “missing” for these purposes, because their employment records are not available.

⁴⁰ The acceptability of the combination of a name and an SSN can vary over time. OCSE reviews the SSN ownership every quarter for the entire sample.

Each quarter, we submitted a match request file to OCSE that contained the names and SSNs for everyone randomized to that date. For those where the SSNs and names aligned, OCSE returned earnings data for the eight most recent quarters in the NDNH, which is lagged by two quarters from the date of the match. This meant that we had up to eight wage reports for each quarter. We used the last version for each quarter within a window. For example, for earnings in the second quarter of 2014, we used reports from the match file for the third quarter of 2016 and discarded the seven earlier sets of earnings data for the second quarter of 2014.

When the earnings data for a quarter contained two or more reports for the same person from the state, we assumed that these reports reflected either different payments by the same employer or payments from different employers. Consistent with the logic discussed in Appendix E, we reviewed quarterly earnings for any values that were clearly impossible, but failing to find any such values, did not discard or top-code⁴¹ any large earnings amounts.

We calculated two outcomes for each quarter: a binary indicator of “any earnings” (yes/no) and the total reported wages for the quarter (\$). The result was two series of 18 measures for each person (employment and earnings for the four quarters before randomization and the 13 quarters after randomization). In addition, we formed a quarterly average earnings for Q12 and Q13 after random assignment (the confirmatory earnings outcome, established to align with the theory of change) and annual averages for Q0-Q3, Q4-Q7, and Q8-Q11.

⁴¹ Meaning values above a threshold are set equal to the threshold.

Appendix D: Comparing NDNH- and Survey-Based Employment and Earnings Estimates

Barnow and Greenberg (2015) review findings from evaluations including both the NDNH and surveys as data sources. Although average survey-reported earnings tend to be higher than average total UI earnings, impact estimates still may be nearly unbiased (Kornfeld and Bloom 1999). In the evaluation of Bridge to Employment, average quarterly earnings agree rather well between the two measurement systems, but correlational analysis shows that there must be considerable measurement noise in one or both. The correlation in person-level quarterly earnings between the two systems at Q12 is just 0.47 for the treatment sample and 0.67 for the control sample.⁴² Earnings from self-employment appear to explain part of the lower correlation on the treatment sample. The difference between self-reported and NDNH-reported earnings has a correlation of 0.12 with self-reported self-employment earnings for the treatment sample, compared to just 0.04 for the control sample.

This section compares estimates of employment and earnings impacts based on NDNH data and survey self-reports.⁴³ It also presents estimates of the impact of Bridge to Employment on self-employment earnings.

The top panel in Exhibit D-1 below shows the degree of agreement of impact estimates for Bridge to Employment derived from the two sources. The estimated impact based on UI records of \$380 for average earnings in Q12 appears to be larger than the estimated impact of \$84 for Q12 based on three-year follow-up survey data. However, the difference between the two estimates is not statistically significant.⁴⁴ We explored whether earnings from self-employment could explain the difference between \$380 and \$84 if we were to treat the difference as real, but earnings from self-employment are too small to explain all of the difference. It could be that the difference is just due to random memory errors by respondents.

Another plausible contributing cause to the discrepancy is differential undercoverage in the NDNH. Barnow and Greenberg (2015) noted that state UI tax databases do not cover federal

⁴² The survey figures convert the available survey measure—earnings in the prior week (calculated as hourly wage multiplied by number of hours worked)—to a calendar quarter-level estimate by multiplying by 13 (the average number of weeks in a quarter).

⁴³ From the follow-up survey, we had a complete history of jobs, with the starting wage and hours for each job as well as the last wage and hours for each job. We combined these to establish weekly earnings for the first and last weeks of a job. We then interpolated to get wages for each intervening month. We then summed weekly wages across jobs for multiple-job holders to get weekly earnings for every week between randomization and interview. Finally, we summarized these to the person-quarter level.

⁴⁴ Assuming a correlation of 0.57 between the two person-level latent effects (the average of the correlations for the two groups), the standard error between the two estimated impacts is \$303, which is larger than the difference between the two impact estimates.

workers; out-of-state records; most workers at small farms, at railroads, at selected nonprofit organizations (particularly churches); and casual or irregular jobs not covered in UI tax databases. Hiding of tip income and income from household employment (such as childcare and cleaning) are additional important sources of undercoverage. In some states, independent contractors, domestic/farm workers, people who work on commission, and other categories of workers are not included. NDNH remedies the undercoverage of federal workers and of out-of-state workers, but not the other causes of undercoverage. If control group members are more likely to find employment of the type undercovered by NDNH, then that could lead to positive bias in the NDNH-based impact. That the NDNH-estimated impact is not statistically significant and that the difference between the two estimated impacts is also not statistically significant, there seems to be no reason to be concerned about this issue, however.

The second panel of Exhibit D-1 shows that NDNH-based employment estimates are slightly higher than survey-based estimates for both treatment group members (76 and 71 percent, respectively) and control group members (74 and 69 percent), leading to somewhat different estimated employment impacts. Most of the difference is probably due to the time frame. The percentage with any earnings over three months is bound to be higher than the percentage employed on a particular day. Nonetheless, the estimated impacts are similar.

Exhibit D-1: Impacts of Bridge to Employment on Earnings and Employment around Follow-up Q12 Based on Wage Records and Self-Reports

Outcome	Treatment	Control	Impact	Standard Error
Quarterly Earnings				
Average NDNH earnings in Q12 (\$)	4,941	4,561	380	258
Self-reported earnings in Q12 (\$)	4,892	4,809	84	363
Self-reported earnings from self-employment in Q12 (\$)	19	5	14	13
Employment				
Average percentage with employer-reported wages in Q12	75.9	74.0	1.9	2.7
Percentage working in the week prior to survey interview	71.2	68.6	2.7	3.7
Sample sizes				
NDNH	506	498		
Survey	346	312		

SOURCES: National Directory of New Hires and PACE three-year follow-up survey.

Note: Self-reported earnings are calculated for the week prior to the survey interview, based on reported work hours and wages, and multiplied by 13 weeks for a quarterly estimate. A majority of survey interviews occurred in the 12th and 13th follow-up quarters.

Statistically significant in a two-tailed test as follows: *** 1 percent level; ** 5 percent level; * 10 percent level.

Appendix E: Treatment of Outliers

We took a conservative approach to outliers, retaining extreme values except where they were clearly impossible. This approach is based on the general difficulty of discriminating between errors and legitimate large values and on the fact that remedies require assumptions about true values that may not be correct.

Trimming observations could easily introduce non-ignorable nonresponse by making nonresponse a function of Y .⁴⁵

Winsorizing observations (also known as “top-coding,” where values above a threshold are set equal to the threshold) could introduce bias if there is a treatment impact but the same threshold is used for treatment and control group members (and there is no reasonable basis for setting different thresholds for the two groups).

Furthermore, evidence suggests that results are generally robust to extreme values. In particular, research by Judkins and Porter (2016) and Lumley, Diehr, Emerson, and Chen (2002) indicates that for the sample sizes available in this evaluation, ordinary least squares inference on the reported data should be robust to outliers.

Outcomes assessed for extreme values included instructional hours (by type of instruction), credits, and NDNH earnings. We found no values that were clearly impossible, and thus retained all reported values in the analysis.

⁴⁵ Trimming by definition creates item nonresponse because the provided response is discarded. If trimming is a function of observed Y , as is standard, and if there is some relationship between observed Y and true Y , then item nonresponse becomes a function of true Y , which is known as “non-ignorable nonresponse.” Because there is no known way to remove bias due to non-ignorable nonresponse, trimming is likely to create uncorrectable biases in estimated treatment effects.

Appendix F: Cost-Benefit Analysis Supplement

This appendix supplies methodological details and reports supplementary findings for the cost-benefit analysis (CBA) in Chapter 6. Methodological details include discussion of assumptions and approximations and descriptions of data sources. Supplemental findings augment, support, and provide additional detail to the higher-level findings reported in Chapter 6.

A cost-benefit analysis plan (Dastrup, Burnett, and Buron 2017) provided an in-depth conceptual overview of the purpose of cost-benefit analyses in the three-year PACE evaluation, along with a plan to conduct CBAs in up to six of the nine PACE programs. That document included discussions of methodological details. That document also sketched several alternative approaches to assumptions and data sources for the analysis. Which alternative would be preferred would depend on the specific program contexts, the pending impact study findings, and data availability for the various programs.

This appendix gives a high-level overview of the approach to the CBA, but does not reprise the conceptual detail provided in the analysis plan. The subsequent sections of this appendix detail the assumptions and data sources used for the CBA of the Bridge to Employment program given the program's unique features (e.g., Individual Training Accounts), the realized impact study findings, and the data available for the analysis.

The main result of the CBA is the *net benefit* of the Bridge to Employment program. Exhibit F-1 below summarizes the findings of the CBA by reproducing the information in Exhibits 6-3, 6-4, and 6-5 from the report. For each component of the analysis, the exhibit reports costs, benefits, and the combined net benefit from each stakeholder group perspective of interest and for society as a whole.

Each cost or benefit component estimate is the difference between the treatment and control group averages for that component. **Net benefit** is then the sum of all benefit components minus the sum of all cost components. In general, the difference is either (1) the treatment group average cost or benefit less the control group average for the same cost or benefit, or (2) a direct estimate from a statistical model of the average difference between treatment and control group of the particular cost or benefit. The CBA used observed study participant and program data wherever possible to measure and monetize inputs and outcomes. Complete data on all benefits and costs were not always available. Then, values for costs or benefits are imputed or approximated using benchmarks from external research or the CBA team's best estimate.

The CBA's focus is whether benefits are greater than costs for society as a whole. However, whether or not the benefits of an intervention justify the costs typically differs when benefits and costs are limited to a given stakeholder's perspective. So the CBA also estimates benefits and costs from the perspectives of treatment group members (relative to the control group), the federal government, California/local government, and the rest of society. These perspective estimates are presented in the columns of Exhibit F-1.

Exhibit F-1: Costs and Benefits of Bridge to Employment, by Perspective

Component	Estimated Net Effect of Bridge to Employment, by Perspective				
	Participants	Government , Federal	Government , State/Local	Rest of Society	Society as a Whole (sum)
Costs (\$)					
A Bridge to Employment services or alternatives available in the community, total	-1,514	3,333	0	0	1,819
<i>Non-ITA elements (assessments, navigators, employment services, supportive assistance, administrative costs)</i>	0	1,819	0	0	1,819
ITAs ^a	-1,514	1,514	0	0	0
B Cost of education or training (including ITA-funded) in three years after random assignment	617 ^b	771	-301	112	1,199
Total Cost (Non-ITA elements + postsecondary education and training)	-897	4,104	-301	112	3,018
Benefits (\$)					
C Net present value of earnings in after random assignment (Q1-Q19)	1,846 ^c	0	0	0	1,846
D Fringe benefits	611	0	0	0	611
E Taxes (federal, state, payroll, sales)	-587	650	78	0	141
F Public assistance (SNAP/WIC, housing assistance) ^d	-316	411	0	0	95
Total Benefits	1,554	1,061	78	0	2,693
Net Benefits = Total Benefits – Total Costs (\$)					
Net Benefit, per Participant Q1-Q19	2,451	-3,043	379	-112	-325

Source: PACE cost data interviews and Bridge to Employment program financial records; PACE 18-month and three-year follow-up surveys; Delta Cost Project Database; Integrated Postsecondary Education Data System; ITA payment records, research team investigation, and approximations of costs of alternative services accessed by the control group. National Directory of New Hires; National Bureau of Economic Research TAXSIM model (Feenberg and Coutts 1993) ; California Department of Tax and Fee Administration; Consumer Expenditure Survey by Income Quintiles (Table 1203); Sources for estimating public assistance listed in Exhibit F-9.

^aPer-participant ITA costs averaged \$2,459 for treatment group members and were approximately \$945 for control group members. These costs are realized by the federal government perspective, and thus contribute approximately \$1,514 of the difference in the cost of education and training. ITAs are used to pay tuition and fees, which are included in the cost of postsecondary education and training.

^bEducation and training costs from the participants' perspective include ITA payments.

^cThis impact estimate has standard error of \$3,058, and an associated *p*-value of .546.

Subsequent sections of this appendix detail the CBA's approach to estimating each of the costs and benefits listed in Exhibit F-1. Costs included in the analysis are for:

- **Bridge to Employment services.** This includes costs for direct services and administrative expenses incurred in operating the Bridge to Employment program, including formal and informal assessments, navigation and case management services, ITAs, supportive services, and employment services (as described in Section 1.1). ITAs are an important feature and primary cost driver of the Bridge to Employment program (see text box in Section 6.1).

Control group members did not have access to the Bridge to Employment program, but they could participate in training, financial assistance, and job search assistance available in the community, notably through the local American Job Centers (AJCs). The estimated cost of access to such services by the control group is used to calculate the added cost of Bridge to Employment.

The federal government funds Bridge to Employment and local AJCs, so costs in this component are allocated to the federal government perspective. In calculating total costs to society as a whole, ITAs are treated as a transfer from the federal government to program participants in the Bridge to Employment services component, represented as a negative cost to participants in Exhibit 6-1. Then the program participants use the ITAs to pay out-of-pocket costs reported in the postsecondary education and training component.⁴⁶

- **Postsecondary education and training.** This includes direct costs associated with all postsecondary education and training (reported in Chapter 3), as measured by institution expenses or inclusive tuition charges (e.g., at for-profit training providers for which expenditure data is not available).⁴⁷ In assessing costs from various perspectives, ITAs are included with other resources (loans and out-of-pocket expenditures) that participants use to pay tuition and fees, resulting in increased cost shown for participants in Exhibit F-1. Federal costs increase as treatment group members attend institutions where federal funding is a higher share of the institutions' non-tuition revenues as compared to those attended by control group members. State and local costs (e.g., appropriations for community colleges) actually decrease, as treatment group members

⁴⁶ The cost of the ITA is included in the final total cost to society as a whole calculation through the education and training component. The cost nets out from the total cost to society perspective in the Bridge to Employment cost component, as it is entered as a transfer from the federal government to participants. This allows the cost to contribute to the total cost to the federal government perspective. The ITA cost nets out from the total cost for the participants perspective, as the negative cost in the Bridge to Employment program component is offset by the positive cost of education and training.

⁴⁷ Direct costs are inclusive of expenditures on instruction, public service, academic support, student services, institutional support, operations and maintenance, scholarships and fellowships, and depreciation, as reported by institutions. The primary indirect cost that is not included here is potential foregone earnings during any training and education funded or induced by Bridge to Employment. Rather than split earnings into during-training and after-training periods, we analyze total earnings over the follow-up period as a component of benefits.

are less likely to attend community colleges because ITAs were more frequently accepted at private institutions.

The bottom panel of Exhibit F-1 shows the primary benefits considered in the CBA. They include:

- **Earnings.** Bridge to Employment's effect on earnings is the main expected benefit. As explained in this chapter's introduction, this CBA includes impacts on total earnings over the first 19 quarters after random assignment. Changes in a number of items that result from increased earnings are included in the CBA:
 - **Fringe benefits.** Increases in earnings and full-time work imply increases in fringe benefits such as health insurance, employer retirement contributions, paid vacation, and sick leave. These benefits represent a value to participants in addition to earnings. The value of fringe benefits is estimated by multiplying observed earnings by external estimates of average benefits as a share of income, adjusted for the proportion of participants that receive benefits.
 - **Taxes.** Increased earnings also generate increases in taxes for treatment group members. The analysis includes estimated amounts for income, payroll, and sales taxes (assuming increased earnings increase taxable purchases). Because taxes represent a transfer from study participants and employers to federal, state, and local governments, the amounts cancel each other out in the society as a whole perspective. Increased employer contributions to payroll taxes based on participants' wages do represent a benefit to society.
 - **Public assistance.** Increases in earnings reduce treatment group members' eligibility for, and receipt of, means-tested public assistance. The loss to participants and benefit to government budgets largely cancel each other out, although there is a small net savings to society from reduced governmental costs for administering programs. Public assistance receipt is estimated by multiplying earnings impacts by external estimates of how increases in income affect benefits, and adjusting for survey-measured rates of benefit receipt by study participants. As detailed below, this CBA considers public assistance items that are included in the PACE surveys, which provide estimates of the incidence of benefit receipt. These items are food assistance (Supplemental Nutrition Assistance Program/SNAP or Special Supplemental Nutrition Program for Women, Infants, and Children/WIC), Temporary Assistance for Needy Families/TANF or other cash public assistance, Unemployment Insurance, housing assistance, and Medicaid (public health insurance).

The first three sections of this appendix provide details of the methodology, with supplementary findings related to the labeled rows of Exhibit F-1.

- Section F.1 details the approach to calculating the cost of the Bridge to Employment program and approximating the cost of control group alternative services. (row A)

- Section F.2 details the approach to estimating the cost of education and training obtained after random assignment and provides supplementary analysis about this cost component. (row B)
- Section F.3 reports the approach to measuring differences in earnings and estimating the fringe benefit, tax, and means-tested public assistance implications of the differences. (rows C, D, and E)

The final two sections are relevant to the entire CBA.

- Section F.5 discusses the uncertainty in the analysis, including how sensitive the findings are to alternative values for assumptions made throughout the analysis.
- Section F.6 catalogues the various data sources used in the CBA.

F.1 Cost of Bridge to Employment Program and Control Group Alternatives

This section describes methods for determining the first cost component of the CBA, the cost of the Bridge to Employment program less the costs of similar services in the community that control group members may have accessed (row A of Exhibit F-1). This section first details the data collection and estimation process used to estimate the average per-participant costs of the Bridge to Employment program. It then turns to the approach for approximating costs of similar services used by the control group.

F.1.1 Cost of the Bridge to Employment Program

To estimate the per-participant cost of Bridge to Employment, the CBA analysis identified all program inputs and expenditures on the inputs. The sum of those expenditures are then divided by the number of treatment group members. To analyze the share of costs by type of input, the CBA also assigned expenditures to cost categories: Individual Training Accounts (ITAs); program activities (formal and informal assessments, navigation and case management services, and employment services); supportive services; and administration and overhead. The resources listed in Exhibit F-2 below are used to develop this estimate of per-participant costs.

Costs by Perspective. The CBA assumed that the entire cost of the Bridge to Employment program is borne by the federal government, as all program expenditures were paid using a federal grant.

Exhibit F-2: Sources Used to Estimate Per-Participant Cost of the Bridge to Employment Program

<i>Resource</i>	<i>CBA Use</i>
San Diego Workforce Partnership program profile (Elkin, Farrell, and Willie 2013) and <i>Implementation and Early Impact Report</i> (Farrell and Martinson 2017)	<ul style="list-style-type: none"> • Preliminary identification of program inputs
Program cost interview with Bridge to Employment program administrators	<ul style="list-style-type: none"> • Confirm comprehensive list of program inputs • Determine that program expenditure reports and ITA tracking log capture costs of all inputs • Assign expenditures to categories
Bridge to Employment program expenditure reports (expenditures by category for fiscal year 2013)	<ul style="list-style-type: none"> • Determine cost of all program inputs except ITAs
Program's ITA tracking log (listing of all ITAs funded by the Bridge to Employment program, including amounts paid and total cost of enrollment)	<ul style="list-style-type: none"> • Determine the cost of ITAs
PACE study enrollment data	<ul style="list-style-type: none"> • Identify study participants enrolled during period covered by expenditure data

F.1.2 Cost of Similar Services Accessed by Control Group Members

The next element needed to calculate the cost of the Bridge to Employment program is the cost of similar services accessed by control group members. The control condition in the evaluation did *not* prohibit access to non-Bridge to Employment services. Control group members did not have access to the Bridge to Employment program, but they could participate in all other training, financial assistance, and job search assistance available in the community.

Sections 3.1.3 and 4.5 of the *Implementation and Early Impact Report* (Farrell and Martinson 2017) document that control group members accessed supportive and employment services, but at a lower rate than treatment group members did.

The CBA approximates the costs of the control groups' use of alternative services in the community. This cost is approximated rather than estimated because, though the PACE survey includes information on the *incidence* of the use of such services, the CBA does not have information on the *quantity* of services that control group members used.

The approximation is the average of two proxies for control group costs: the first is based on Bridge to Employment costs, which are scaled down based on the survey-reported incidence of control group services used relative to the treatment group and the amount of assistance available to each group. The second is an estimate of the per-participant cost of basic Workforce Investment Act services from the recent literature.

The first proxy scales down the observed per-participant costs of Bridge to Employment. Exhibit F-3 below develops the approximation multipliers we use for this scaling. For ITAs, the CBA bases the multiplier on (1) how often program participants may have used ITAs, based on the incidence of education or training; (2) reported incidence of financial assistance receipt on the follow-up survey; and (3) a lower cap on ITA amounts available through AJCs as compared to the Bridge to Employment program.

Exhibit F-3: Reported Differences in Level of Services Accessed by Control Group Members

Outcome	Evidence from Prior PACE Research			Conclusions Drawn for CBA Approximation	
	Treatment Group	Control Group	Control Group as % of Treatment Group	Approximation Multiplier	Value of Multiplier (%)
ITAs					
Received education/training since random assignment (%): In any subject/field ^a	75.2	57.8	77	Opportunity to use ITAs	77
Received financial assistance at first place of instruction (%): Grants/scholarships ^b	62.2	49.8	80	Reported use of grants/scholarships	80
Financial assistance for training: ITA cap ^c	\$7,000 or \$10,000	\$5,000	80 ^d	Amount of ITAs	80
				Combined ITA multiplier	49 ^f
Program Activities					
Received assistance from any organization since random assignment (%) ^b					
Career counseling	32.5	24.4	75	Program activities multiplier	67 ^g
Job search or placement	35.8	18.9	53		
Supportive Assistance					
Received assistance from any organization since random assignment (%): Help arranging supports for school/work/family ^b	25.0	11.3	45	Supportive assistance multiplier	41 ^f
Supports: Supportive services cap ^a	\$1,000	\$500	90 ^e		

Source: *Implementation and Early Impact Report* (Farrell and Martinson 2017); Bridge to Employment program ITA logs; PACE three-year follow-up survey.

^a Exhibit 4-8, *Implementation and Early Impact Report*.

^b Exhibit 4-9, *Implementation and Early Impact Report*.

^c Exhibit 3-1, *Implementation and Early Impact Report*.

^d Approximately 70 percent of Bridge to Employment ITAs are below \$5,000. The research team assigned 80 percent for this multiplier as an approximation.

^e Qualitative data collection for the implementation report and amounts indicated in itemized expense summaries suggest that supportive services assistance in Bridge to Employment rarely exceeds \$500.

^f Calculated as products; that is, $77\% \times 80\% \times 80\% = 49\%$; and $45\% \times 90\% = 41\%$.

^g The research team selected 67 percent as a rough approximation between 75 percent and 53 percent.

For program activities costs, the CBA selects an approximation multiplier of 67 percent, based on the average difference in the incidence of receipt of career counseling and job search or placement reported by treatment group and control group members in the 18-month follow-up survey.⁴⁸ For supportive services costs, the multiplier of 41 percent is based on the incidence of receiving help arranging supports for school/work/family on the survey and on a lower cap on supportive services.

To approximate control group costs per participant of similar services obtained elsewhere, we applied these multipliers to the respective category of Bridge to Employment program costs per participant. We then summed across the approximated amounts for each category and added the 6 percent administrative cost (same as observed for Bridge to Employment) to determine the first proxy of control group costs.

The second proxy is based on an estimate from recent literature on the cost of job training at AJCs. Fortson et al. (2017) report an estimate of the average cost of “core services” provided by 28 Workforce Investment Act programs around the country. The CBA uses these core services costs (reported in their Table VIII.1) to approximate costs of services that control group members may have accessed.⁴⁹ The CBA proxies ITAs, program activities, and supportive assistance costs by multiplying the analogous \$579, \$506, and \$58 cost estimates from Fortson et al. or similar services by 1.17, which is the regional price parity value published by the Bureau of Economic Analysis for the San Diego metropolitan statistical area in 2013.⁵⁰ The same 6 percent administrative cost rate observed for Bridge to Employment costs is used for the proxy estimates.

Exhibit F-4 below summarizes the proxies and final approximation. For reference and comparison, the first column of the exhibit reproduces the estimated costs per treatment group member of the Bridge to Employment program. The second column shows the first proxy amounts, which total \$2,793. The research team assessed that this proxy is likely too high. Control group members likely received less-intensive services than did participants in Bridge to Employment (not accounted for by this proxy) in addition to the observed lower incidence of any service receipt. This conclusion is based both on descriptions of the availability and use of services by control group members reported in the program profile (Elkin, Farrell, and Willie 2013) and early impact report (Farrell and Martinson 2017) and on feedback from the Bridge to

⁴⁸ With this multiplier, we are assuming that differences in the incidence of any use of services approximate differences in quantity of services used, and that unit costs of alternative services are the same as for Bridge to Employment services.

⁴⁹ Random assignment in the PACE evaluation was conducted after applicants had been in contact with a navigator organization and had been introduced to the Bridge to Employment program and agreed to participate in the study. Thus, they were interested in obtaining services that might improve their employment situation, including the career counseling and training that the program provided.

⁵⁰ The federal Bureau of Economic Analysis’s Regional Price Parities by Metro Area All Items (MARPP) index. Retrieved from <https://www.bea.gov/data/prices-inflation/regional-price-parities-state-and-metro-area> on 10/2/2019.

Employment study liaisons who had conducted on-site interviews with program staff familiar with the program and the availability and use of alternate resources.

The third column of Exhibit F-4 shows the resulting second proxy amounts, which total to \$1,420. This second approach results in about 50 percent lower cost than the first. The final approximation of the cost of similar services accessed by control group members is the simple average of the two proxies (fourth column). The Chapter 6 CBA reports and uses this approximation.

Exhibit F-4: Summary of Proxies Used to Approximate Cost per Control Group Member of Similar Services

Service Type	Cost per Treatment Group Member— Bridge to Employment	Cost per Control Group Member		
		Proxy 1— Approximation Multiplier × Bridge to Employment Cost	Proxy 2— AJC Average Adjusted to San Diego	Final Approximation
ITA Element (\$)				
ITAs	2,459	1,213	676	945
Non-ITA Elements (\$)				
Assessments, navigators/case management, and employment services	2,598	1,732	592	1,162
Supportive services	269	109	66	88
Administrative cost	344	198	86	142
Total	3,211	2,039	744	1,392
Total Costs (\$)				
Total	5,670	3,252	1,420	2,337

Source: PACE cost data interviews and Bridge to Employment program financial records; PACE 18-month and three-year follow-up surveys; Fortson et al. (2017).

Costs by Perspective. The CBA assumes that any alternative services accessed in the community are also funded by federal grants.

F.1.3 Cost of the Bridge to Employment Program

The cost of the Bridge to Employment program is then the estimated cost per treatment group member of the program minus the approximated cost per control group member of alternative services accessed in the community. The CBA separates the non-ITA and ITA amounts as subtotals of this cost in row A of Exhibit F-1. ITAs are entered as a positive cost to the federal government and a negative cost to participants. Then, in the next row, education and training costs from the participants perspective include the portion covered by ITA payments. These costs are detailed in Section F.2 below. This is an efficient approach to including these costs without double-counting. Specifically, the CBA does not link specific ITA use to follow-up survey responses on education and training spells used to estimate costs. Rather, the CBA estimates costs of the education and training reported in the survey whether the cost is funded by an ITA or not.

F.2 Cost of Education or Training

This section considers the second cost component of the CBA, the estimated effect of assignment to the Bridge to Employment treatment group on costs of education and training enrollment after random assignment (row B of Exhibit F-1). The section first reports methods and data used for estimation. It then provides some supplementary analysis that buttresses the assertion that all costs of the Bridge to Employment are realized within the first three years after random assignment.

The cost of the Bridge to Employment intervention should include the costs of all education and training obtained after random assignment. The intervention is expected to affect enrollment in education and training, which results in a change in costs to society. Education and training (and associated costs) will increase if, for example, education and training obtained as an initial result of the Bridge to Employment program is the first of multiple stages along a new or accelerated career trajectory.⁵¹ Alternatively, Bridge to Employment services could initiate a successful career for treatment group members that replaces less directed education and training and thus lowers costs to society of education and training. Thus, included in this category are both costs of the initial college or career training courses that were funded using Bridge to Employment program ITAs for treatment group members and costs associated with additional education that may be a follow-on result of program participation.

F.2.1 Methods and Data

The CBA estimates the cost of education or training as the product of a *quantity* measure of units of education and training received and a *unit cost* of the education and training. The quantity of education obtained is measured from individual-level information from the three-year follow-up survey. We estimate unit cost from external institution-level estimates.

To determine the quantity of education or training received, the CBA builds on the Chapter 3 analysis of impacts on postsecondary education or training. Exhibit 3-1 in Chapter 3 reports a quantity measure of education or training obtained: full-time-equivalent (FTE) months enrolled at any school. The CBA directly adopts this measure of quantity, which is based on person-level education spells constructed from the three-year follow-up survey. The CBA adds the reported institution attended to each spell to create a dataset of FTE months at each institution attended by academic year for each survey respondent.

The CBA determines the unit cost of each FTE month as an institution's total cost of expenditure per-FTE month of enrollment. This value is calculated or imputed for each institution where study participants reported enrollment using data from the U.S. Department of Education's Integrated Postsecondary Education Data System (IPEDS) supplemented with additional study data. The primary source of data for unit costs is the Delta Cost Project Database (DCPD), an extract of educational institution finances from IPEDS (Hurlburt, Peek,

⁵¹ This concept of career pathways is common to all PACE programs, and implies that Bridge to Employment participants may obtain additional education and training later in their career pathway because of their participation in the PACE program.

and Sun 2017).⁵² Exhibit F-5 details the variables we use from the DCPD along with a few additional IPEDS variables. All dollar-denominated variables are expressed in terms of FTEs and adjusted for inflation to 2013 dollars using public access code downloaded with the DCPD.⁵³

Exhibit F-5: DCPD/IPEDS Variables Used in the CBA

DCPD/IPEDS Variable	Variable Description from DCPD/IPEDS Documentation	Variable Use in CBA
Total education and general expenditures—current year total (eandg01_fte_cpi)	Includes all core operating expenditures, including sponsored research, but excluding auxiliary enterprises. This variable was originally reported in IPEDS, but for recent years it is calculated by summing expenditures on instruction, research, public service, academic support, student services, institutional support, operations and maintenance, and scholarships and fellowships.	Primary component of total per-participant costs.
Depreciation—current year total (depreciation01_fte_cpi)	Allocation or distribution of the cost of capital assets, less any salvage value, to expenses over the estimated useful life of the asset in a systematic and rational manner. Depreciation for the year is the amount of the allocation or distribution for the year involved. This field is used if the institution has not allocated all depreciation to other functions.	Added into total per-participant costs where reported.
Expenditures for other non-operating—current year total (othernon01_fte_cpi)	Other non-operating expenses and deductions. Total expense is the sum of all expenses incurred other than interest that are not classified as operating expenses.	Added into total per-participant costs where reported.
Non-educational share of expenses (noneducation_share)	Research and public service portion of spending on instruction, student services, research, and public service.	Used to prorate total per-participant costs to education share of expenses only. Imputed as state-level sector by highest-degree-offered mean where missing.
Sector of institution (sector)	One of nine institutional categories resulting from dividing the universe according to control and level. Categories are public 4-year or above, private nonprofit 4-year or above, private for-profit 4-year or above, public 2-year, private nonprofit 2-year, private for-profit 2-year, public less-than-2-year, private nonprofit less-than-2-year, private for-profit less-than-2-year.	Used to define “similar” institutions for imputing values using means.

⁵² We developed an FTE month count for each spell in the follow-up survey responses based on responses about whether students were attending full-time or part-time. FTE months were calculated for DCPD data by dividing the IPEDS variable on annualized full-time-equivalent undergraduate enrollment by 12. That variable is defined in IPEDS documentation as follows: “For institutions with a semester, trimester, or 4-1-4 plan, the number of FTE undergraduates is the sum of undergraduate credit hours divided by 30 and contact hours divided by 900. For institutions with a quarter plan, undergraduate credit hours divided by 45 and contact hours divided by 900. For institutions with continuous enrollment over a 12-month period, undergraduate credit hours were divided by 30 and contact hours were divided by 900.”

⁵³ The code uses a CPI-U inflation adjustment variable that is included with the data download.

DCPD/IPEDS Variable	Variable Description from DCPD/IPEDS Documentation	Variable Use in CBA
Highest degree offered (hdegofr1)	Eight institutional categories defining highest degree offered by institution. Available values are non-degree-granting, doctor's degree–research/scholarship and professional practice, doctor's degree–research/scholarship, doctor's degree–professional practice, doctor's degree–other, master's degree, bachelor's degree, and associate degree.	Used to define “similar” institutions for imputing values using means.
Net tuition and fees revenue (nettuition01)	Amount of money the institution takes in from students after institutional grant aid is provided (this is not the same as the net tuition number available in IPEDS, which is net of all discounts and allowances applied to tuition and fees).	Numerator when calculating tuition and fees share of cost for perspective analysis.
Net tuition directly from students (net_student_tuition)	Net tuition revenue coming directly from students (not including Pell, federal, state, and local grants).	Numerator when calculating student out-of-pocket costs share for perspective analysis.
Pell grants (grant01_fte_cpi)	Gross amount of Pell grants disbursed or otherwise made available to recipients by the institution.	Numerator when calculating the Pell and other federal grant awards cost share for perspective analysis.
Other federal grants (grant02_fte_cpi)	Expenditures for scholarships and fellowships, excluding Pell grants, which were funded from federal government agencies. This includes Supplemental Educational Opportunity Grants (SEOG) and State Student Incentive Grants (SSIG), but not loans or College Work Study Program.	
Revenue from federal appropriations, grants, and contracts (federal10)	The total amount of revenue coming from federal appropriations, grants, and contracts.	Less Pell and other federal grants is numerator for calculating federal non-grant cost share for perspective analysis.
Revenue from state and local appropriations (state_local_app)	The total amount of revenue from state and local appropriations.	Combined to determine the numerator for the state and local cost share for perspective analysis.
Revenue from state and local grants and contracts (state_local_grant_contract)	The total amount of revenue from state and local grants and contracts. Grants by state government include expenditures for scholarships and fellowships that were funded by the state.	
State grants (grant03) and local grants (grant04)	Grants by local government are for scholarships and fellowships that were funded by local government.	
Revenue from affiliated entities, private gifts, grants, and contracts; investment returns; and endowment earnings (priv_invest_endow)	The total amount of revenue coming from affiliated entities, private gifts, grants and contracts, investment returns, and endowment earnings. Endowment earnings stopped being reported to IPEDS in 1997 for FASB-reporting institutions and 2002 for GASB-reporting institutions.	Numerator when calculating the other cost share for the perspective analysis.

Key: FASB = Financial Accounting Standards Board. GASB = Governmental Accounting Standards Board. IPEDS = Integrated Postsecondary Education Data System

Source: Integrated Postsecondary Education Data System and Delta Cost Project Database data documentation.

Study participants attended a number of institutions that do not submit data to IPEDS. For these institutions, the CBA approximates costs per FTE month enrolled using alternative study data sources and secondary research. Notably, such institutions include those that most readily accepted Bridge to Employment ITAs: for-profit career training organizations that provide

healthcare training and certifications and public less-than-two-year continuing education schools. Data collected from the Bridge to Employment program included logs tracking the program's issuance of ITAs. The logs include a *full cost of training* field, which we use to estimate the cost of an FTE month at these institutions. The research team augmented this information by consulting these institutions' websites for information on tuition and fees.⁵⁴ The research team also reviewed financial statements and collected enrollment information for public continuing education schools.

The CBA research team conducted such institution-specific research for all institutions that did not contribute financial data to IPEDS and had non-trivial enrollment by study participants. For institutions with few enrollment spells by study participants, the cost of an FTE month is imputed as the state average of institutions for which there is data in the same sector (e.g., public two-year) with the same highest degree grantee (e.g., associate degree), weighted by FTE months attended by study participants at the other institutions.

The CBA uses a five-year average, 2011-2015, of annual DCPD unit cost estimates to estimate net costs of education and training. DCPD data are only available through the 2014-2015 school year, whereas education spells recorded in the PACE three-year follow-up survey responses extend into the 2016-2017 school year. So the analysis must impute unit cost data for 2015-2016 and 2016-2017 in order to calculate total costs. Rather than applying observed cost estimates in earlier years and imputed unit costs in later years, the CBA uses a five-year average of inflation-adjusted, observed costs for all periods. Trends in unit costs during the study period are not anticipated to contribute to *net* costs, and so the CBA does not attempt to model or approximate trends in costs for any institutions.

To approximate individual-level costs of all education and training, the CBA multiplies each individual quantity estimate (FTE month spell) by the relevant institution's unit cost estimate (five-year average cost per FTE month estimate). All of an individual's observed education and training spells are then summed. Because education and training costs occur over multiple years, the analysis must account for the time value of money by calculating a net present value (NPV) at random assignment—the sum of costs over time, discounted to reflect the time elapsed since random assignment. The CBA discounts the spell costs at a 3 percent annual rate and sums all discounted spell costs. Thus we assumed that \$1 at random assignment is valued equivalently to \$1.03 a year later. See Dastrup et al. (2017) for further motivation. As noted above, the DCPD variables are already adjusted for inflation.

The CBA uses this individual-level cost of education and training as the outcome in a statistical impact model that estimates the net cost of education and training. Specifically, the individual-level cost estimate is treated as an outcome, and the CBA estimates an impact on that outcome as is done for other outcomes in the impact study. The CBA uses the same impact model used to estimate impacts on postsecondary education and training in Chapter 3. This approach provides an estimate of the net cost of education and training that results from the Bridge to Employment program as the difference between average treatment group member costs and

⁵⁴ As they appeared at the time of the study using internet archive websites where possible.

average control group member costs. The resulting estimated impact on education and training costs is reported in Exhibit 6-2 and incorporated into Exhibit F-1.

Costs by Perspective. To assign costs of education and training to different perspectives, the CBA relies on two types of estimates. First, the CBA estimates out-of-pocket expenses for study participants. Second, it estimates the share of an institution's revenues that are from various sources that correspond to each stakeholder perspective. Overall revenue shares are used to estimate cost shares for stakeholder perspectives (except participants themselves) because institution educational expenditures (realized costs) are typically not directly linked to specific revenue sources. That is, the CBA assumes an institution's revenues are fungible across all educational expenditures.⁵⁵

The CBA estimates out-of-pocket expenses of education and training for study participants using the same data sources that we use to approximate unit costs. It first calculates the average net student out-of-pocket tuition and fee amount (which is often negative, meaning students are remitted Pell and other grants). This amount will include students' ITA payments. The following details of education and training providers and study participant characteristics informed the estimation.

For student out-of-pocket expenses, the CBA assumes the following (for both treatment and control groups).

- Study participants pay the approximated out-of-pocket tuition and fees less the amount of grants received.
 - Grants received are the average federal (primarily Pell) and state grant amounts received by first-time full-time students receiving grants reported in the DCPD. Grants received are zero at institutions that are not eligible to provide grants (i.e., for-profit less-than-two-year career training institutions).
 - As discussed in the short-term report, study participants are almost certain to qualify to attend community colleges in California with all enrollment fees waived.⁵⁶ So we approximate that participants pay zero out-of-pocket tuition and fees to enroll at any community colleges in California. The CBA thus assumes that study participants receive the average grant amounts remitted to them as (a negative cost of attendance) at these institutions.
 - At all other institutions (e.g., four-year public and private institutions and for-profit career training institutions), out-of-pocket tuition and fees are approximated as

⁵⁵ The analysis does, however, assume that other students' out-of-pocket tuition and fees support their own costs of attendance, rather than subsidize PACE participant's costs.

⁵⁶ This is currently called the California College Promise Grant, formerly called the BOG Fee Waiver. Waived fees (i.e., fees that are not paid by students who may or may not have enrolled without a fee waiver) do not represent a realized cost of education to the State of California. Rather, we use grants and appropriations that provide the actual revenue used for educational expenditures to determine costs to the state.

the average student out-of-pocket costs, either observed in the DCPD or approximated using the same approach for unit costs described above.

- ITA amounts from the Bridge to Employment program logs are not linked to student enrollments in the three-year follow-up survey. The CBA assumes that the calculated average net student out-of-pocket tuition and fee amount includes payments made by students using ITAs, and it includes a negative cost to participants of ITAs in the Bridge to Employment component.

The costs of education and training from the remaining perspectives are estimated using institution-specific revenue shares calculated from DCPD variables, determined through research team investigation, or imputed based on the other similar institutions that study participants attend. Specifically, the CBA calculates average per-participant costs of education and training for each subgroup perspective as follows:

- For the federal government perspective, total costs are multiplied by the share of institutions' non-tuition and fee revenues that are from federal sources. These are primarily student-level grants (Pell and others), but also include other federal appropriations.
- For the state and local government perspective, total costs are multiplied by the share of institutions' non-tuition and fee revenue that is from state and local appropriations and grants.
- For the rest of society perspective, total costs are multiplied by the revenue share from other sources (e.g., endowment revenue and private donations).

F.2.2 Supplementary Analysis

This section provides additional information and analysis for education and training costs. First, it provides examples of institution-level unit cost estimates. Then, it shows estimates of the cost of education and training by quarter since random assignment.

Cost Characteristics of Most-Attended Institutions

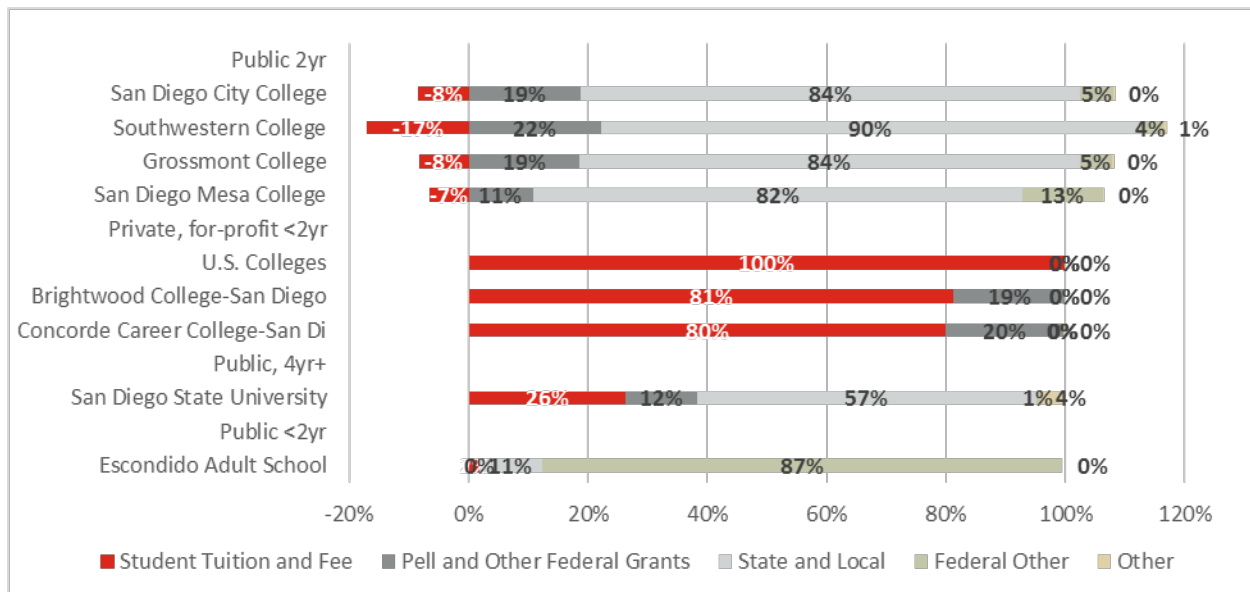
As discussed in Section 6.2, the initial *Implementation and Early Impact Report* (Farrell and Martinson 2017) found that treatment group members were less likely to attend a two-year college than control group members were. Instead, they were more likely to attend a private, non-degree-granting school or adult high school or adult learning courses. Exhibit F-6 below provides specific examples of the difference in mix and cost of programs attended. The exhibit lists the share of all FTE months attended (across about 100 total institutions) for the treatment and control groups for nine selected institutions. The institutions include the five institutions most attended by either treatment group or control group members. These nine examples cover roughly half of all FTE months attended by each group. The exhibit also lists the estimate of the institution's per-FTE monthly cost.

Exhibit F-6: Per-FTE Monthly Total Costs at Most-Attended Institutions

Institution	Share of All FTE Months Attended in Q1-Q13 (%)		Estimate of Institution's Per-FTE Monthly Total Cost of Instruction (\$)
	Treatment Group	Control Group	
Public 2-Year			
San Diego City College	8.5	5.1	1,069
Southwestern College	4.7	12.1	921
Grossmont College	3.1	10.3	1,065
San Diego Mesa College	2.0	4.9	1,654
Subtotal (weighted by FTE months)	18.3	32.4	1,098
Private For-Profit Less-than-2-Year			
U.S. Colleges ^a	7.9	1.8	3,163
Brightwood College	7.4	4.7	897
Concorde Career College	5.2	6.8	895
Subtotal (weighted average)	20.5	13.3	1,553
Public 4-Year-Plus			
San Diego State University	6.5	4.1	1,403
Public Less-Than-2-Year			
Escondido Adult School	5.9	2.9	1,000

Source: PACE three-year follow-up survey; Integrated Postsecondary Education Data System; Delta Cost Project Database; ITA payment records; research team analysis.

^a U.S. Colleges is the name of a single for-profit, less-than-two-year healthcare career training institution with multiple locations in Southern California.

Exhibit F-7: Per-FTE Monthly Total Costs at Most-Attended Institutions


Source: PACE three-year follow-up survey; Integrated Postsecondary Education Data System; Delta Cost Project Database; ITA payment records; research team analysis.

Note: Shares for student out-of-pocket revenue can be negative because institution's total payments to students can exceed tuition and fees received from students.

Exhibit F-7 above reports the break out of revenue across these sources for the same most-attended institutions included in Exhibit F-6. Annual shares for student out-of-pocket revenue can be negative because an institution's total payments to students (e.g., as disbursements of Pell and other grants) can exceed tuition and fees received from students.

Quarter-by-Quarter Analysis of Education and Training Costs

The CBA estimates education and training from the three-year follow-up survey. As a result, any impact on education training past three years after random assignment is not included. This section presents supplementary findings addressing the likelihood of differences in education and training cost between treatment and control group members that persist beyond the three-year follow-up survey.

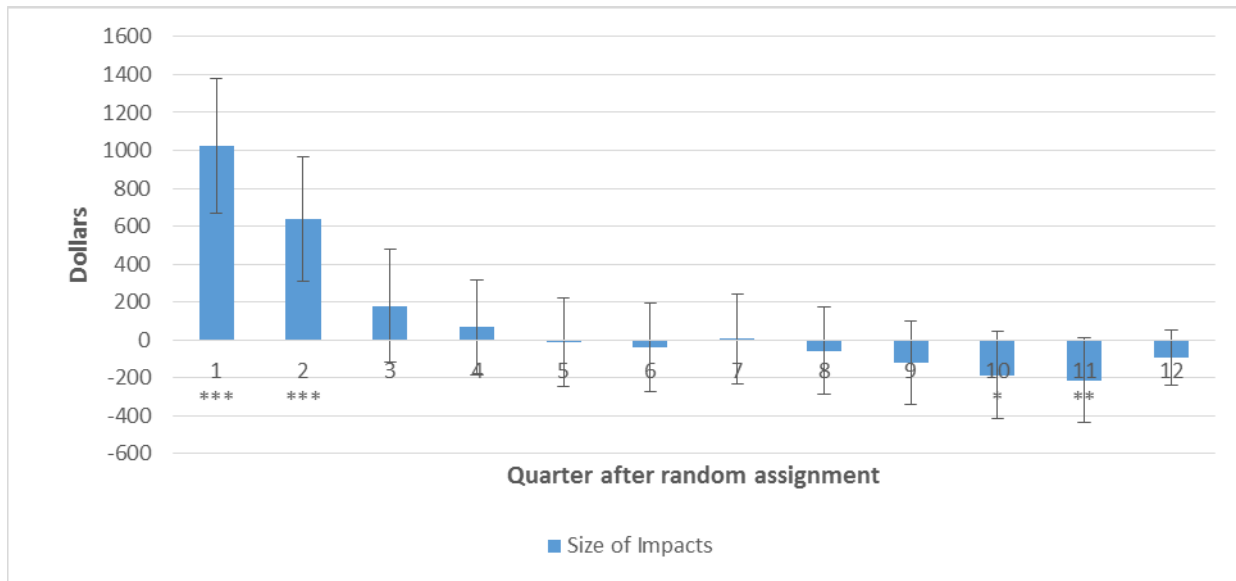
Chapter 6 argues that essentially all net costs of the Bridge to Employment program were incurred in those first three years. Three facts support a conclusion that costs are not understated due to the three-year data window for education and training because all costs of the Bridge to Employment program were incurred in those three years.

First, all study members have completed their engagement with the Bridge to Employment program itself.

Second, treatment group members typically used ITAs in training that began within the first quarter after random assignment (about 70 percent of the time), and almost always within a year (97 percent). The length of ITA-funded training was typically a few months, although some programs (about 5 percent) extended over a year or more. Thus, almost all education and training directly funded by Bridge to Employment ITAs was complete by the second year after random assignment (well within the data window). Our education and cost estimates include this period plus an additional third year, so no further costs can be directly attributed to the program.

Third, a quarter-by-quarter analysis shows that differences in education and training costs for treatment and control group members occurred primarily in the first three quarters after random assignment. Exhibit F-8 reports education and training cost impacts by quarter. These impacts are quarterly net cost estimates. These differences in costs between the treatment group and control group were largest (and statistically significant) in Q1 and Q2. In Q5 and beyond, net costs are negative and small, and only statistically significant in Q10 and Q11. The CBA has only three years of data on the costs of education and training (fewer than the 19 quarters of data for earnings), but it appears that any meaningful differences in education and training costs between the treatment and control groups occurred by Q12.

Exhibit F-8: Education and Training Costs Impacts by Quarter



Source: PACE three-year follow-up survey; Integrated Postsecondary Education Data System; Delta Cost Project Database; ITA payment records; research team analysis.

Statistical significance levels based on differences between research groups: *** 1 percent level; ** 5 percent level; * 10 percent level.

F.3 Earnings Impacts, Fringe Benefits, Taxes, and Means-tested Assistance

Earnings impacts reported in the CBA (row C of Exhibit F-1) are estimated with the same impact model as earnings impacts in Chapter 4, except that individual-level earnings are discounted to the time of random assignment to account for inflation and the time value of money. The CBA calculated the NPV at the time of random assignment as the sum of all discounted earnings. The nominal discount rate of 5 percent aligns with the 3 percent used for the inflation-adjusted education and training costs under an assumption that inflation is approximately 2 percent.

Increases in earnings and full-time work imply increases in fringe benefits such as health insurance, employer retirement contributions, paid vacation, and sick leave (row D, Exhibit F-1). Estimates in this analysis multiply earnings gains by external estimates of average fringe benefit value. This approach follows Schaberg and Greenberg's (2020) CBA of WorkAdvance. To develop assumptions about appropriate multipliers and the value of fringe benefits as a percentage of earnings, the CBA research team consulted the Employer Costs for Employee Compensation (ECCC), Compensation Percentiles, from the National Compensation Survey, produced by the U.S. Bureau of Labor Statistics and national averages of fringe employment benefit receipt from Solis and Galvin (2012).

Following Schaberg and Greenberg, the ratio of the value of total fringe benefits to total wages for private industry workers reported in the ECCC from 2016 to 2018 averaged 42.5 percent. Bridge to Employment participants' median wages were comparable to the median wages over this period in the ECCC, so we assume that the 42.5 percent value of benefits applies to treatment group participants earnings impacts. A comparison of figures reported in Solis and Gavin (2012) to averages for Bridge to Employment treatment group participants indicates that

treatment group participants receive fringe benefits at a lower rate than all workers (similarly to WorkAdvance participants in Schaberg and Greenberg, who receive benefits at approximately 60 percent of the prevailing rates). Additional analysis of PACE three-year follow-up survey responses shows that fringe benefit receipt rates are below national averages (with no statistically significant differences between the treatment and control group), with rates of benefit receipt that average roughly 77.9 percent of the national average. So, to estimate increases in the value of fringe benefits, the CBA multiplies earnings impacts by $0.425 \times 0.779 = 0.331$.

The CBA estimated the tax implications of earnings changes (row E of Exhibit F-1) using Internet TAXSIM (v27), a cloud-based program of the National Bureau of Economic Research.⁵⁷ TAXSIM returns a listing of the applicable federal and state tax liabilities, including both income and payroll taxes, for submitted taxpayer profiles. The CBA calculates earnings impacts using NDNH data. The research team was not allowed to transmit individual-level profiles that include NDNH earnings to the TAXSIM program in order to calculate individual tax liabilities. Instead, tax liabilities were calculated for a variety of household structures at the observed average earnings levels for the whole sample. The resulting variety of tax liabilities were then averaged by the incidence of each household structure in the study sample.

Specifically, 12 profiles were submitted for each of five tax years (2013-2017): filing status (joint or single), by number of dependents (zero, one, two), at each year's average treatment and control group level of earnings based on the sums of quarterly impact estimates. For joint filing households, spouses are assumed have \$10,000 in wages or salary. The analysis assumes one dependent is under age 13 and the second (if any) is under age 18. The six profiles are then recombined for each treatment condition using the family study proportions reported in the *Implementation and Early Impact Report* (see Farrell and Martinson 2017, Exhibit A-2). Specifically, the proportions are 46.4 percent single with no dependents, 14.1 percent each for single with one and two dependents (28.2 percent combined), 12.0 percent joint with no dependents, and 6.7 percent each for joint with one and two dependents. The top panel of Exhibit F-9 lists the elements of the approach to estimating changes in taxes associated with observed changes in earnings.

⁵⁷ <https://users.nber.org/~taxsim/taxsim27/>

Exhibit F-9: Estimating Marginal Effective Taxes Associated with Earnings Impacts

Element Estimated	Sources	Estimate Details
Income Taxes (federal, state, and payroll)		
Effective tax amounts	NBER Internet TAXSIM and analysis of NDNH earnings data	Tax profiles using 6 household types and average treatment and control group member earnings by year
Study participants' household type shares	<i>Implementation and Early Impact Report</i> Appendix Exhibit A-2	Single no dependents: 46.4% Single one dependent: 14.1% ^a Single two dependents: 14.1% Joint no dependents: 12.0% Joint one dependent: 6.7% ^a Joint two dependents: 6.7%
Effective Sales Tax		
Average share of income spent on taxable purchases	Consumer Expenditure Survey by Income Quintiles (Table 1203) CDTFA ^b	1/3 of income spent on sales taxable goods and services
Sales tax rate	CDTFA	8.00%
<i>Implied Multiplier</i>		2.67%
Marginal Effective Tax – SNAP/WIC		
Share of study participants affected	Analysis of baseline and three-year survey	42.6%
Estimate of marginal effective tax rate for food assistance	Marginal rate effect for those affected (<i>weighted by earnings</i>)	11.2%
<i>Implied Multiplier</i>		4.8%
Marginal Effective Tax – TANF or other cash public assistance		
Share of study participants affected	Analysis of baseline and three-year survey	14.0%
Estimate of marginal effective tax rate for cash public assistance	Calculated based on Hanson and Andrews (2009)	48.1%
<i>Implied Multiplier</i>		6.8%
Change in receipt of Unemployment Insurance		
Change in households receiving unemployment or workers compensation	Analysis of three-year survey	1.1 percentage points
Value of Unemployment Insurance	Estimate from Vroman (2018)	\$4,200 per instance
<i>Implied Savings</i>		\$46
Marginal Effective Tax – Housing		
Share of study participants affected	Analysis of baseline and three-year survey	10%
Marginal effective tax rate for housing assistance	Most federal housing assistance programs require recipients to pay 30% of income in rent	30%
<i>Implied Multiplier</i>		3.1%

Key: CDTFA = California Department of Tax and Fee Administration. NBER = National Bureau of Economic Research. SNAP = Supplemental Nutrition Assistance Program. WIC = Special Supplemental Nutrition Program for Women, Infants, and Children.

^a The actual number of dependents is not available in PACE data. Households with children are assumed to have one or two dependents in equal proportion.

^b CDTFA (2019): Sales and Use Taxes: Exemptions and Exclusions, www.cdtfa.ca.gov/formspubs/pub61.pdf; Historical Tax Rates in California Cities & Counties, www.cdtfa.ca.gov/taxes-and-fees/archive-rates.htm.

^c Joint Center for Housing Studies (2017).

Note: The 2015 tax year is used as a mid-point of the earnings year, where applicable.

The second panel of Exhibit F-9 shows the approach to accounting for sales tax. The first resource is a tabulation of average expenditure shares by category for low-income households from the Consumer Expenditure Survey reported by the Bureau of Labor Statistics.⁵⁸ These shares are cross-tabulated by a list of categories of purchases subject to sales tax published by the California government, and determine that approximately one third of low-income households' income is spent on items subject to sales tax. The sales tax rate in San Diego is approximately 8 percent. Together, this implies a sales tax offset on income gains of 2.67 percent.

The CBA also calculates marginal effective taxes for each year for changes in sales taxes, and public assistance measured in the 36-month follow-up survey. These include food assistance (SNAP and WIC), TANF or other cash public assistance, unemployment insurance, and housing assistance (row E of Exhibit F-1). The bottom section of Exhibit F-9 provides the details to the estimation approach for each of these. The CBA estimates average marginal effective taxes of 2.7 percent for sales taxes, 5.3 percent for SNAP/WIC benefits, and 7.5 percent for housing assistance.

The marginal effective tax is also relevant for how possible future earnings gains may update CBA conclusions. The impact model estimates a \$587 increase in earnings in the final year of available follow-up data (Q16-19), with treatment group average earnings of \$22,396 and control group average earnings of \$21,809. The CBA estimates that study participants would face a combined effective marginal tax on this increase of about \$289, including both federal and state taxes and means-tested assistance. This includes increases of \$118 in net federal income taxes, \$10 in state income taxes, \$45 in the employee portion of FICA taxes, and \$15 in sales taxes; and losses of \$101 in public assistance. This amount may be understated because we may not observe all means-tested assistance (such as subsidized childcare).

Earnings, Taxes, and Means-tested Assistance by Perspective

Assigning earnings and associated taxes and means-tested assistance to the various perspectives is a relatively straightforward exercise of noting which perspective is receiving money and, for elements that are transfers, which perspective is paying it. Transfers are amounts that flow from one perspective to another and so net out for society as a whole.

Exhibit F-10 summarizes the assignment of net costs and benefits to each perspective; it includes earnings, taxes, and means-tested assistance in its second panel. The benefit of earnings accrues to participants. Taxes are transfers from participants to state/local government, meaning higher tax amounts are a net negative benefit for participants and a net positive benefit for government. A small transfer of the employer portion of payroll taxes from the rest of society to the federal government is also included. Means-tested assistance are transfers from governments to participants.

⁵⁸ <https://www.bls.gov/cex/tables.htm>.

Exhibit F-10: Summary of Net Costs and Net Benefits by Perspective

Component	Description of Net Effect of Bridge to Employment, by Perspective				
	Treatment Group (relative to control)	Government, Federal	Government, State/Local	Rest of Society	Society as a Whole (sum)
Net Costs					
Net cost of the non-ITA components of Bridge to Employment program or alternative services available in the community ^a		Net cost			Net cost
Cost of education or training (including ITA-funded) in three years after random assignment	Net benefit	Net cost	Net cost	Net cost	Net cost
Total Net Cost	Net benefit	Net cost	Net cost	Net cost	Net cost
Net Benefits					
NPV of earnings in Q1-Q19 after random assignment	Net benefit				Net benefit
Net taxes (federal, state, payroll, including credits)	Net cost	Net benefit	Net benefit	Net cost	Transfers offset
Public assistance (SNAP/WIC, housing assistance) ^a	Net cost	Net benefit			Transfers offset
Total Net Benefits	Net benefit	Net benefit	Net benefit	Net cost	Net benefit
Summary					
Total Net Benefits – Total Net Costs, per Participant Q1-Q19	Net benefit	Net cost	Net cost	Net cost	Net cost

Key: NPV = net present value. SNAP = Supplemental Nutrition Assistance Program. WIC = Special Supplemental Nutrition Program for Women, Infants, and Children.

^a To avoid confusion in terminology in this chapter, we use the term “public assistance” when referring to what other chapters call “public benefits” (e.g., TANF, SNAP, WIC, Medicaid, etc.).

F.4 Uncertainty in Components of the Cost-Benefit Analysis

Estimated NPVs of PACE programs based on all measured costs and benefits are subject to three types of uncertainty:

- (1) Sample variability;
- (2) Measurement error in a single observation of Bridge to Employment costs; and
- (3) A multiplicity of options for elements that cannot be estimated from observed data but must instead be assumed from estimates available in the CBA literature.

This section discusses each of these sources of uncertainty. After describing each type of uncertainty, we discuss how sensitive the CBA findings likely are to the particular uncertainty. In brief, the uncertainty associated with earnings impact estimates is large enough such that definitive CBA findings are not possible, and other sources of uncertainty do not materially affect this conclusion. In addition to the uncertainty discussed in this section, Section 6.4 discusses

two additional sources of uncertainty—intangible costs and benefits that are not monetized and items that could affect the CBA that are not included in the analysis.

F.4.1 Sample Variability and Measurement Error in Participant-level Data

Sample variability and measurement error in participant-level data⁵⁹ cause virtually all impact estimates and other parameters calculated using statistical analysis to be subject to some uncertainty. This is true even for parameters found to be statistically significant. Standard errors associated with each estimate provide a measure of the extent of this uncertainty. Larger standard errors indicate greater parameter uncertainty.

Confidence intervals are one way of expressing these standard errors. The CBA reports 90 percent confidence intervals in Chapter 6 for the two elements of the CBA for which estimates are based on participant-level data that use a statistical model of impacts. The 90 percent confidence interval for the estimate of the net costs of education and training spans from $-\$534$ to $+\$2,930$. The 90 percent confidence interval for the estimate of the net benefit of the NPV of quarterly earnings after random assignment spans from $-\$3,188$ to $+\$6,881$.⁶⁰

Section 6.6 discusses how sensitive the CBA findings are to the uncertainty of these estimates. That discussion concludes that given the imprecision in the estimates of earnings and education and training impacts, definitive CBA conclusions are not possible.

Uncertainty also exists for the elements in the perspective analysis calculated using the same data on costs of education and training used to calculate total costs of education and training. These elements are program participants' estimated out-of-pocket costs and the share of revenue that institutions derive from a variety of sources.⁶¹

F.4.2 Measurement Error in Single Program and Site Estimates

The CBA estimates Bridge to Employment program costs based on cost data collected for a single site. Inaccuracies in observed values result in measurement error for this cost estimate. Because costs are incurred (and observed) at the program level for a single program, no standard error or comparable characterization of the resulting uncertainty in a site-level point estimate exists. However, because actual ITA costs and program expenses for treatment group members are observed for a relatively time limited program, it is unlikely that the size of error in the measurement of program costs is materially relative to the standard errors of the estimates

⁵⁹ Sample variability and measurement error both result in chance variation in outcomes not due to the Bridge to Employment program that is in part a result of error in measuring the outcome.

⁶⁰ These two estimates are based on the same sample of individuals and thus are likely correlated. We nevertheless do not attempt to characterize any correlation in uncertainty between these outcomes. Doing so would require joint estimation of the impacts, which is not straightforward given our general approach to estimation of impacts.

⁶¹ Unlike overall costs, we estimate statistically significant differences in the share revenues coming from student tuition and fees, state and local appropriations, and Pell and other federal grants. This analysis is correlational because it is weighted by participant FTE months attended at each institution and is thus conditional on any attendance.

measured with individual-level data. This means that measurement error of Bridge to Employment program costs is unlikely to affect the CBA conclusions.

Additionally, the CBA calculates the net cost of the program by subtracting an approximation of the cost of control group member use of similar alternative services. The next section discusses the uncertainty associated with this approximation as uncertainty in assumed parameters.

F.4.3 Error in Assumed Parameters

Many parameters in the CBA must be assumed rather than estimated from available data. These assumed parameters are documented in the prior sections of this appendix or in Chapter 6. Such assumed parameter values characterize the extent of similar career advancement services accessed by control group members, elements of the calculation of the marginal effective taxes in the perspective analysis, and the choice of a discount rate to account for the time value of money.

Could errors in the approximation of control group access to services in the community similar to those provided Bridge to Employment participants materially affect the CBA conclusions? The total cost of the Bridge to Employment program per treatment group member is \$5,670. The CBA approximates that alternative access to services by the control group cost \$2,106. Exhibit F-3 in Section F.2.2 provides an alternative proxy 1 estimate of about \$2,800, which the research team assesses to be too high. At the other extreme, if control group members received no services whatsoever (contrary to related evidence in follow-up survey responses), the CBA would be understating the net cost by \$2,106. This amount is a similar order of magnitude as the estimated margin of error of the cost of education or training, and substantially smaller than half the span of the 90 percent confidence interval of the earnings estimate.

Using a lower-bound or higher-bound cost estimate in place of the final approximation would not change the final CBA conclusions. That is, at the lower bound, the best estimate of total net benefits minus total net costs is more negative, but uncertainty associated with the earnings net benefits would still result in the conclusion that the estimate is not precise enough to draw a definitive conclusion. And using the higher-bound estimate of proxy 1, the best estimate negative would be smaller, but still non-negligible, and we would still not draw a definitive conclusion that net benefits are higher, or lower, than net costs.

Because the calculation of marginal effective taxes is based on fractions of observed earnings, error in the approach to these calculations is some fraction of the margin of error of the earnings impact estimates. Additionally, this estimation process largely identifies transfers that affect the perspective analysis only, but not the main cost-benefit result from the perspective of society as a whole.

The CBA uses a 3 percent real discount rate to calculate the NPV at random assignment of costs of education and training undergone in the subsequent three years. A real discount rate is used because the DCPD data amounts are already adjusted for inflation to 2013 dollars. Two (2) percentage points are added to this discount rate for the earnings analysis to account for inflation, resulting in a 5 percent nominal rate. It is standard practice in CBAs to recalculate NVPs using higher and lower alternative discount rates, such as 2 and 6 percent (real)

(Boardman et al. 2011).⁶² The CBA recalculated earnings impacts using 3 and 8 percent nominal discount rates. Results are presented in Exhibit F-11. The impact estimate is about \$150 higher when using the lower discount rate, and \$200 lower when using the higher discount rate. Thus, the choice of discount rate is not material for the CBA conclusions.

Exhibit F-11: Net Present Value of Quarterly Earnings after Random Assignment

Outcome	Discount Rate	Treatment Group	Control Group	Impact (Difference)	Standard Error	Relative Impact	p-Value
Net present value of total earnings after random assignment (Q1-Q19)	3%	\$73,333	\$71,332	+\$2,000	\$3,246	2.7%	.538
	5%	\$69,144	\$67,299	+\$1,846	\$3,058	2.7%	.546
	8%	\$63,507	\$61,869	+\$1,638	\$2,806	2.6%	.560
Sample size		506	498				

Source: National Directory of New Hires.

In principle, it is possible to assess the likely combined effect of these many sources of uncertainty by conducting a Monte Carlo analysis. For such a Monte Carlo analysis, the total net benefits minus total net costs CBA conclusion would be calculated a very large number of times (more than 10,000) with the parameters subject to substantive uncertainty all drawn with each calculation from distributions of their probable values. In the Monte Carlo framework envisioned in Dastrup et al. (2017), NPV values for all parameters that could materially alter CBA conclusions would be replaced simultaneously with each calculation. However, the uncertainty associated with earnings impacts (and to a lesser extent education and training costs) dominates others sources of uncertainty in the analysis. As a result, the characterization of uncertainty for this variable presented in Section 6.6 is an adequate characterization of the uncertainty for the entire CBA.

F.5 Data Sources

This section catalogues the information and data sources used to develop the estimates reported in Chapter 6.

F.5.1 Program Profile and Implementation and Early Impact Report

Profiles of each of the nine PACE programs, including Bridge to Employment, (Elkin et al. 2013), were compiled as part of the PACE evaluation implementation study.⁶³ The Bridge to Employment profile gives a high-level overview of the program; details program goals, target population, and structure; and describes the program's career pathways components. The information on program structure and career pathways components was used to conduct background research that identified program inputs. This research was completed in preparation for the cost data collection interviews we conducted for the CBA.

⁶² This recalculation is less relevant in our CBA because we examine earnings for 19 quarters after random assignment, a time frame over which choice of discount rates has relatively small consequences.

⁶³ Program profiles are available at <http://www.acf.hhs.gov/opre/research/project/pathways-for-advancing-careers-and-education>.

The *Implementation and Early Impact Report* (Farrell and Martinson 2017) included program descriptions and analysis of 18-month follow-up survey data. The research team consulted the report to develop a proxy for control group member access to similar services.

F.5.2 Qualitative and ITA Log Data from the PACE Implementation Study

The PACE program profiles were based on site visits in 2012 and 2013 that included interviews with program leadership and staff, review of documents related to the program, and observation of program activities. Site visit teams also had monthly monitoring calls in which they discussed study enrollment and program implementation. The research staff who conducted the site visits and maintained contact with the programs kept and organized notes documenting these visits, which were used to produce the profiles.

The research team found these notes useful in preparing for cost data collection interviews. In particular, these notes sometimes included additional detail on program components and structure that had not been included in the program profiles. For Bridge to Employment, the person who had conducted the program's site visit remained as a research team member, so the CBA research team met with that person to review the program's structure and context.

The CBA research team also reviewed logs created as administrative records tracking all ITAs issued by the Bridge to Employment program. The logs were used to estimate per-participant ITA receipt and the per-FTE month cost at institutions where ITAs were used.

F.5.3 Cost Data Collection Interviews and San Diego Workforce Partnership Expenditure Reports

In 2015, we conducted interviews with San Diego Workforce Partnership staff to gather information on Bridge to Employment program operations and costs. Program directors and financial officers participated in a one- to two-hour phone interview. The research team used this interview and subsequent follow-ups to determine a comprehensive list of program inputs and associated costs. The program also provided annual, audited expenditure reports to estimate the cost of the program.

F.5.4 Follow-up Surveys

The CBA used follow-up survey data (from 18-month and three-year surveys) to determine the participant education and training spells and to inform approximations of control group services costs.

F.5.5 DCPD and Other IPEDS data

The Delta Cost Project Database is a longitudinal database derived from the U.S. Department of Education's Integrated Postsecondary Education Data System. The CBA calculated five-year averages of the expenditure, revenue, enrollment, and student aid variables used in the analysis, from the 2011-2015 school years. IPEDS data added additional information to determine key institution characteristics (such as type of institution and undergraduate share). The database includes information on more than 6,000 public, private not-for-profit, and private for-profit institutions, including revenues, sources of revenues, spending, and total operating expenditures (Hurlburt et al. 2017).

F.5.6 NDNH Wage Data

Derived from state Unemployment Insurance records, the National Directory of New Hires contains quarterly employment/earnings data for all covered workers. The CBA used NDNH data for the 19 quarters after random assignment—the longest time period available among the exploratory outcomes for earnings.

F.5.7 TAXSIM Simulation Output

Internet TAXSIM (v27) is a cloud-based program that “calculates federal and state income tax liabilities from typical survey data” available and documented via the National Bureau of Economic Research website (Feenberg and Coutts 1993). Internet TAXSIM takes as input a data file of taxpayer profiles, including tax year, state, marital status, number of dependents of various ages, and wage and salary income. For each submitted profile, TAXSIM returns a listing of the applicable federal and state tax liabilities, including both income and payroll taxes. The CBA uses the TAXSIM analysis to estimate the tax implications of earnings impacts.

Appendix References

- Barnow, B. S., and D. Greenberg. 2015. "Do Estimated Impacts on Earnings Depend on the Source of the Data Used to Measure Them? Evidence from Previous Social Experiments." *Evaluation Review* 39 (2): 179-228.
- Beicht, Ursula, and Michael Friedrich. 2008. "Anlage und Methode der BIBB-Übergangsstudie." In *Ausbildungschancen und Verbleib von Schulabsolventen*, edited by Ursula Beicht, Michael Friedrich, and Joachim Gerd Ulrich, 79-99. Bielefeld, Germany: W. Bertelsmann.
- Betz, N. E., and K. M. Taylor. 2001. *Manual for the Career Decision Self-Efficacy Scale and CDMSE—Short Form*. Columbus, OH: The Ohio State University.
- Boardman, Anthony, David Greenberg, Aidan Vining, and David Weimer. 2011. *Cost-Benefit Analysis: Concepts and Practice*, 4th ed. Upper Saddle River, NJ: Pearson Higher Education.
- Bühlmann, P., and S. van de Geer. 2011. *Statistics for High-Dimensional Data*. Berlin, Heidelberg, Germany: Springer.
- California Department of Tax and Fee Administration. 2019. *Sales and Use Taxes: Exemptions and Exclusions. California Revenue and Taxation Code Part 1, Division 2*. Retrieved from <https://www.cdtfa.ca.gov/formspubs/pub61.pdf>.
- Ciolino, Jody D., Hannah L. Palac, Amy Yang, Mireya Vaca, and Hayley M. Belli. 2019. "Ideal vs. Real: A Systematic Review on Handling Covariates in Randomized Controlled Trials." *BMC Med Res Methodol* 19: 136. Doi:10.1186/s12874-019-0787-8
- Cohen, S., R. Kamarck, and R. Mermelstein. 1983. "A Global Measure of Perceived Stress." *Journal of Health and Social Behavior* 24 (4): 385-396.
- Congressional Budget Office (CBO). 2015. *Effective Marginal Tax Rates for Low- and Moderate-Income Workers in 2016*. <https://www.cbo.gov/sites/default/files/114th-congress-2015-2016/reports/50923-marginaltaxrates.pdf>.
- Cutrona, C., and D. Russell. 1987. "The Provisions of Social Relationships and Adaptation to Stress." *Advances in Personal Relationships*, 1.
- Dastrup, Samuel, Kimberly Burnett, and Larry Buron. 2017. *Career Pathways Intermediate Outcomes Study: Plan for Cost-Benefit Analyses* (OPRE Report 2017-68). Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. <https://www.acf.hhs.gov/opre/resource/career-pathways-intermediate-outcomes-study-plan-for-cost-benefit-analyses>.
- Deke, J., and H. Chiang. 2017. "The WWC Attrition Standard: Sensitivity to Assumption and Opportunities for Refining and Adapting to New Contexts." *Evaluation Review* 41: 130-154.

- Deville, J. C., and C. E. Särndal. 1992. "Calibration Estimation in Survey Sampling." *Journal of the American Statistical Association* 87: 376-382.
- Dundar, A., & Shapiro, D. 2016. *The National Student Clearinghouse as an integral part of the national postsecondary data infrastructure*. Retrieved from the National Student Clearinghouse Research Center website: <https://nscresearchcenter.org/wp-content/uploads/NSC-as-an-Integral-Part-of-the-National-Postsecondary-Data-Infrastructure.pdf>
- Dynarski, S. M., Hemelt, S. W., & Hyman, J. M. 2015. "The Missing Manual: Using National Student Clearinghouse Data to Track Postsecondary Outcomes." *Educational Evaluation and Policy Analysis*, 37(1s), 53S–79S.
- Elkin, Sam, Mary Farrell, and Jessica Willie. 2013. *Pathways for Advancing Careers and Education Career Pathways Program Profile: San Diego Workforce Partnership's Bridge to Employment in the Healthcare Industry* (OPRE Report 2013-20). Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. <http://www.career-pathways.org/career-pathways-san-diego-workforce-partnerships-bridge-to-employment-in-the-healthcare-industry-program-pdf-14-pp-5-1-mb-august-2013/>.
- Farrell, M., and K. Martinson. 2017. *The San Diego County Bridge to Employment in the Healthcare Industry Program: Implementation and Early Impact Report* (OPRE Report 2017-41). Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. <https://www.acf.hhs.gov/opre/resource/san-diego-county-bridge-employment-healthcare-industry-program-implementation-early-impact-report>.
- Feenberg, Daniel Richard, and Elizabeth Coutts. 1993. "An Introduction to the TAXSIM Model." *Journal of Policy Analysis and Management* 12 (1): 189-194. <http://www.nber.org/taxsim/>.
- Fein, D. J. 2012. *Career Pathways as a Framework for Program Design and Evaluation: A Working Paper from the Pathways for Advancing Careers and Education (PACE) Project* (OPRE Report 2012-30). Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. https://www.acf.hhs.gov/sites/default/files/opre/cp_as_a_framework_final_508b.pdf.
- Folsom, R. E. 1991. "Exponential and Logistics Weight Adjustments for Sampling and Nonresponse Error Reduction. In *Proceedings of the American Statistical Association, Social Statistics Section*, 197-202. Alexandria, VA: American Statistical Association.
- Folsom, R. E., and A. C. Singh. 2000. "The Generalized Exponential Model for Sampling Weight Calibration for Extreme Values, Nonresponse, and Post-Stratification." In *Proceedings of the American Statistical Association, Section on Survey Research Methods*, 598-603. Alexandria, VA: American Statistical Association.

- Folsom, R. E., and M. Witt. 1994. "Testing a New Attrition Nonresponse Adjustment Method for SIPP." In *Proceedings of the American Statistical Association, Section on Survey Research Methods*, 428-433. Alexandria, VA: American Statistical Association.
- Fortson, Kenneth, Dana Rotz, Paul Burkander, Annalisa Mastri, Peter Schochet, Linda Rosenberg, Sheena McConnel, and Ronald D'Amico. 2017. *Providing Public Workforce Services to Job Seekers: 30-month Impact Findings on the WIA Adult and Dislocated Worker Programs*. Washington, DC: Employment and Training Administration, U.S. Department of Labor. <https://clear.dol.gov/study/providing-public-workforce-services-job-seekers-30-month-impact-findings-wia-adult-and-1>.
- Goldrick-Rab, S., and K. Sorensen. 2010. "Unmarried Parents in College." *Future of Children* 20 (2): 179-203.
- Hanson, Kenneth, and Margaret S. Andrews. 2009. *State variations in the food stamp benefit reduction rate for earnings: Cross-program effects from TANF and SSI cash assistance*. ERS Report Summary, Economic Research Services, U.S. Department of Agriculture.
- Hendra, Richard, and Aaron Hill. 2018. "Rethinking Response Rates: New Evidence of Little Relationship between Survey Response Rates and Nonresponse Bias." *Evaluation Review*. <https://doi.org/10.1177/0193841X18807719>.
- Holland, Paul W. (1986). "Statistics and Causal Inference". *J. Amer. Statist. Assoc.* 81 (396): 945–960. doi:10.1080/01621459.1986.10478354
- Hoven, M. R. 2012. "Investigating the Relationship between Perceived Social Support and Parent Self-Efficacy in Parents of Preschool-Aged Children." Master's Thesis. University of British Columbia. https://circle.ubc.ca/bitstream/handle/2429/43343/ubc_2012_fall_hoven_michaelyn.pdf?sequence=3. Last accessed 8/28/2015.
- Hurlburt, Steve, Audrey Peek, and Jie Sun. 2017. *Delta Cost Project Database 1987–2015: Data File Documentation*. Washington, DC: Delta Cost Project at American Institutes for Research. <http://www.deltacostproject.org/>.
- Izrael, David, David C. Hoaglin, and Michael P. Battaglia. 2000. "A SAS Macro for Balancing a Weighted Sample." In *Proceedings of the Twenty-Fifth Annual SAS Users Group International Conference*, Paper 275. Cary, NC: SAS Users Group International.
- Joint Center for Housing Studies of Harvard University. 2017. *America's Rental Housing 2017*. Cambridge, MA: Joint Center for Housing Studies of Harvard University, Harvard Graduate School of Design, Harvard Kennedy School.
- Judkins, D., D. Morganstein, P. Zador, A. Piesse, B. Barrett, and P. Mukhopadhyay. 2007. "Variable Selection and Raking in Propensity Scoring." *Statistics in Medicine* 26: 1022-1033.
- Judkins, David R., and Kristin E. Porter. 2016. "Robustness of Ordinary Least Squares in Randomized Clinical Trials." *Statistics in Medicine* 35 (11): 1763-1773.

- Judkins, David, David Fein, and Larry Buron. 2018. *Analysis Plan for the PACE Intermediate (Three-Year) Follow-up Study* (OPRE Report 2018-95). Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
https://www.acf.hhs.gov/sites/default/files/opre/pace_three_yearanalysisplan_mainreport_508.pdf.
- Judkins, David. 2019. "Covariate Selection in Small Randomized Studies." Presentation at the Joint Statistical Meetings, Denver 2019.
- Koch, G. G., C. M. Tangen, J.-W. Jung, and I. A. Amara. 1998. "Issues for Covariance Analysis of Dichotomous and Ordered Categorical Data from Randomized Clinical Trials and Non-Parametric Strategies for Addressing Them." *Statistics in Medicine* 17: 1863-1892.
- Kornfeld, R., and H. Bloom. 1999. "Measuring Program Impact on Earnings and Employment: Do Unemployment Insurance Wage Reports from Employers Agree with Surveys of Individuals?" *Journal of Labor Economics* 17 (1): 168-197.
- Le, H., A. Casillas, S. Robbins, and R. Langley. 2005. "Motivational and Skills, Social, and Self-Management Predictors of College Outcomes: Constructing the Student Readiness Inventory." *Educational and Psychological Measurement* 65 (3): 482-508.
- Lin, W. 2013. "Agnostic Notes on Regression Adjustments to Experimental Data: Reexamining Freedman's Critique." *The Annals of Applied Statistics* 7, 295-318.
https://projecteuclid.org/download/pdfview_1/euclid.aoas/1365527200.
- Lumley, T., P. Diehr, S. Emerson, and L. Chen. 2002. "The Importance of the Normality Assumption in Large Public Health Data Sets." *Annual Review of Public Health* 23: 151-169.
- Peterson, C. H., A. Casillas, and S. B. Robbins. 2006. "The Student Readiness Inventory and the Big Five: Examining Social Desirability and College Academic Performance." *Personality and Individual Difference* 41 (4): 663-673.
- Research Triangle Institute. 2012. *SUDAAN Language Manual, Volumes 1 and 2, Release 11*. Research Triangle Park, NC: Author.
- Rubin, Donald B. 1987. *Multiple Imputation for Nonresponse in Surveys*. New York, NY: Wiley.
- Schaberg, K., and D. H. Greenberg. 2020. *Long-term Effects of a Sectoral Advancement Strategy: Costs, Benefits, and Impacts from the WorkAdvance Demonstration*. New York: MDRC. <https://www.mdrc.org/publication/long-term-effects-sectoral-advancement-strategy>.
- Solis, Hilda L., and John M. Galvin. 2012. "Labor force characteristics by race and ethnicity, 2011." *Bureau of Labor Statistics*.
- Stumpf, S. A., S. M. Colarelli, and K. Hartman. 1983. "Development of the Career Exploration Survey (CES)." *Journal of Vocational Behavior* 22 (2): 191-226.

- Tukey, John W. 1991. "Use of Many Covariates in Clinical Trials." *International Statistical Review* 59 (2):123-137.
- Vroman, Wayne. 2018. *Unemployment Insurance Benefits; Performance Since the Great Recession*. Urban Institute, Washington DC.
- Walker, Joan M. T., Andrew S. Wilkins, James R. Dallaire, Howard M. Sandler, and Kathleen V. Hoover-Dempsey. 2005. "Parental Involvement: Model Revision through Scale Development." *The Elementary School Journal* 106 (2): 85-104.
- Williams, R. L., and R. E. Folsom. 1981. "Weighted Hotdeck Imputation of Medical Expenditures Based on a Record Check Subsample." In *Proceedings of the American Statistical Association, Section on Survey Research Methods*, 406-411. Alexandria, VA: American Statistical Association.