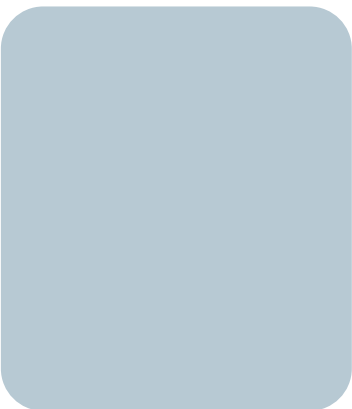
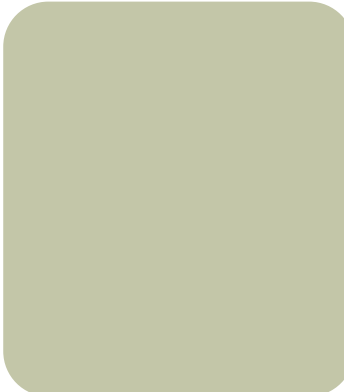


Bridging the Opportunity Divide for Low-Income Youth: Implementation and Early Impacts of the Year Up Program

Appendices

May 2018

OPRE Report #2018-65



PACE
Pathways for Advancing
Careers and Education

Bridging the Opportunity Divide for Low-Income Youth: Implementation and Early Impacts of the Year Up Program Pathways for Advancing Careers and Education (PACE)

OPRE Report #2018-65

May 2018

David Fein and Jill Hamadyk, Abt Associates

Submitted to:

Nicole Constance

Federal Project Officer

Office of Planning, Research, and Evaluation

Administration for Children and Families

U.S. Department of Health and Human Services

Contract No. HHSP2332007913YC

Project Director: Karen Gardiner

Abt Associates Inc.

4550 Montgomery Ave.

Bethesda, MD 20814

This report is in the public domain. Permission to reproduce is not necessary. Suggested citation: Fein, D. and Hamadyk, J. (2018). *Bridging the Opportunity Divide for Low-Income Youth: Implementation and Early Impacts of the Year Up Program*, OPRE Report #2018-65, Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. To contact the first author, e-mail David_Fein@Abtassoc.com.

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 <http://www.acf.hhs.gov/programs/opre/index.html>.



Contents

Appendix A: Baseline Data and Adjustments.....	1
Appendix B: Unemployment Insurance Wage Data	11
Appendix C: National Student Clearinghouse Data	14
Appendix D: 18-Month Follow-up Survey Data.....	16
Appendix E: Comparing NDNH- and Survey-Based Employment and Earnings Estimates.....	28
Appendix F: Treatment of Outliers.....	30
Appendix G: Quarterly Impact Detail for Chapter 6 Exhibits	31

List of Exhibits

Exhibit A-1: Operationalization of Baseline Measures Used as Covariates in Regression-Adjusted Impact Estimates	2
Exhibit A-2: Coefficients for Baseline Characteristics as Predictors of Average Quarterly Earnings for Quarters 6 and 7 after Random Assignment: Year Up Control Group Members	8
Exhibit A-3: Comparison of Selected Impact Estimates With and Without Adjustment for Baseline Imbalances.....	10
Exhibit C-1: NSC School-level Cooperation Rates by School Control and Level from 2012 through 2016	14
Exhibit D-1: Details on Specifications for Survey-Based Outcomes in Chapter 5.....	17
Exhibit D-2: Details on Specifications for Survey-Based Outcomes in Chapter 6.....	19
Exhibit D-3: Baseline Balance on Full Sample, Unweighted Respondent Sample, and Weighted Respondent Sample	24
Exhibit D-4: Comparison of Selected Impact Estimates for the Unweighted and Weighted Survey Samples	27
Exhibit E-1: Impacts on Earnings and Employment around Follow-up Quarters 6-7 Based on Wage Records and Self-Reports	29
Exhibit G-1: Impacts on Earnings and Employment in Successive Follow-up Quarters	31
Exhibit G-2: Impacts on College Enrollment and Cumulative Full-time Equivalent Months of Enrollment in Successive Follow-up Quarters	32

Appendix A: Baseline Data and Adjustments

This Appendix describes specifications for baseline covariates and summarizes the approach to missing values (Section A.1). As summarized in Chapter 3, of 28 characteristics examined, distributions for treatment and control group members differed for only one characteristic—a difference most likely due to chance. This appendix also explains how the study’s approach to estimating impacts controls for these covariates (Section A.2).

A.1 Details on Baseline Covariates

Exhibit A.1 details 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 under four 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).

Imputation for missing covariate values used SUDAAN/IMPUTE: a weighted hot-deck imputation procedure (Research Triangle Institute 2012). Imputation entailed a single computer run on the combined sample from all nine PACE sites. With this process, each missing value was replaced with an observed response from a similar case. Within specified strata, cases with missing values were random-matched to cases with reported values; the reported value was then copied over to the case where the value was missing. The strata represented a cross-classification of: treatment-control status, site, NSC-reported enrollment status (some or none),¹ NSC-reported credential award (some or none), and number of months of NSC-reported enrollment.²

¹ The National Student Clearinghouse (NSC) has information on monthly enrollment and many credentials for 96% of college students. <https://nscresearchcenter.org/workingwithourdata/>.

² In instances where this level of matching was too restrictive because no matched case with a reported value was found, then the procedure was re-run matching only on treatment status and NSC-reported enrollment status.

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: 20 or under 21 to 24 25 or older*	BIF: B2_dob RABIT: R_RA_Date_Assigned
Female	Binary variable 1 if female 0 otherwise	BIF: B7
Race-Ethnicity	Categorical measure: Hispanic, any race Black, non-Hispanic (includes those checking multiple races) White, non-Hispanic* (excludes those checking multiple races) Other, non-Hispanic (residual)	BIF: B9
Living Arrangements	Categorical measure: Not living with spouse/partner or children Not living with spouse/partner, living with children* Living with spouse/partner, not living with children Living with spouse/partner and children (Only biological and adopted children of randomized participant considered here. Step children, grandchildren, younger siblings, and other children not considered.)	BIF: B13
Living with Parents	Binary variable 1 if living with own parent(s) 0 otherwise (Presence of parents of spouse not considered.)	BIF: B13
Educational Background		
At Least One Parent with Some College	Binary variable: 1 if either parent attended college 0 otherwise	BIF: B21
High School Grades	Categorical measure: Mostly A's Mostly B's Mostly C's or below*	BIF: B23
Educational Attainment	Categorical measure: No college* Under 1 year's college 1 year+ of college Associate's degree or higher	BIF: B17

(Continued)

Variable Description	Operationalization Details	Data Source(s) Survey Instrument & Survey Item Number
Career Knowledge		
Career Knowledge Index	<p>A seven-item scale based on a review of existing survey instruments and literature. The first two scale items (a-b) below adapted from the Career Decision Self-Efficacy-Short Form (Betz and Taylor, 2001). Items d-f adapted from Career Exploration Survey. Two items (c and g) were new. Average percentage of respondents answering "strongly agree" to in questions about confidence in different areas of career knowledge. Missing if four or more of seven responses blank.</p> <p>a. You know how to accurately assess your abilities and challenges?</p> <p>b. You know how to make a plan that will help achieve your goals for the next five years?</p> <p>c. You know how to get help from staff and teachers with any issues that might arise at school?</p> <p>d. You know the type of job that is best for you?</p> <p>e. You know the type of organization you want to work for?</p> <p>f. You know the occupation you want to enter?</p> <p>g. You know the kind of education and training program that is best for you?</p>	SAQ: S13
Psycho-Social Indices		
Academic Discipline ³	Average of ten items (scale ranging 1-6) after reversing responses to negatively-phrased items. Missing if seven or more of ten responses blank.	SAQ: S11a
Training Commitment ⁴	Average of ten items (scale ranging 1-6) after reversing responses to negatively-phrased items. Missing if seven or more of ten responses blank.	SAQ: S11b
Academic Confidence ⁵	Average of twelve items (scale ranging 1-6) after reversing responses to negatively-phrased items. Missing if nine or more of twelve responses blank.	SAQ: S11d
Emotional Stability ⁶	Average of twelve items (scale ranging 1-6) after reversing responses to negatively-phrased items. Missing if nine or more of twelve responses blank.	SAQ: S11e
Stress ⁷	Average of four items (scale ranging 1-5) after reversing responses to negatively-phrased items. Missing if three or more of four responses blank.	SAQ: S14

(Continued)

³ 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).

⁴ 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).

⁵ 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).

⁶ 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).

⁷ Cohen, et al. (1983).

Variable Description	Operationalization Details	Data Source(s) Survey Instrument & Survey Item Number
Resource Constraints (Financial)		
Family Income Last Year ⁸	Categorical measure: Less than \$15,000 \$15,000-29,999 \$30,000 or more*	BIF: B27
Received WIC or 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 if 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 otherwise	SAQ: S8, S9
Resource Constraints (Time)		
Current Work Hours	Categorical measure: 0-19* 20-34 35+	BIF: B24
Expected Work Hours in Next Few Months	Categorical measure for covariate: 0-19* 20-34 35+	SAQ: S2
Plan to attend school only part-time if admitted to Year Up	Binary variable: 1 if yes 0 if no	SAQ: S1
Resource Constraints (Financial)		
Life Challenges Index	A new scale created adapted from a longer instrument by Kessler, et al. (1998). Average of six items of frequency of situations that interfered with school, work, job search, or family responsibilities. The response categories ranged from 1='never' to 5='very often'. Missing if four or more of six responses blank. <ul style="list-style-type: none"> • Child care arrangements • Transportation • Alcohol or drug use • An illness or health condition • Arguments with a family member • Physical threats/violence from a family member 	SAQ: S15

Data source abbreviations: RABIT (Random Assignment and Baseline Information Tool), BIF (Basic Information Form), SAQ (Self-Administered Questionnaire). * = category omitted in creating binary (dummy) variables for regression-adjustment models.

⁸ Used as a covariate in analyses of NSC- and survey-based outcomes but inadvertently omitted from covariate list for NDNH-based outcomes.

A.2 Regression Adjustment

This section describes the regression adjustment approach used to improve precision and minimize effects of sampling error on impact point estimates.

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

$$Y_i = X_i\beta + \delta T_i + e_i, \tag{A.1}$$

where Y_i is the outcome, T_i is a 0/1 dummy variable indicating treatment group membership, X_i is a row vector of baseline covariates, β is the vector of parameters indicating the influence of each covariate on the outcome, δ is the effect of treatment, and e_i is an error term. 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 covariates (Lin, 2013) even when the outcomes are not normally distributed (Judkins and Porter, 2016). Estimates of the treatment effect are “asymptotically unbiased” and for adequately large sample sizes, under most conditions, $\text{var}(\hat{\delta}) \approx (1 - R^2) \text{var}(\bar{y}_T - \bar{y}_C)$, where R^2 is proportion of the variance in Y_i that can be explained by X_i , in equation A.2 below.

Preparatory work with PACE data showed that the method can perform poorly when the number of baseline covariates is relatively large compared to the number of observations. Specifically, when the ratio n/p is not very large, it can happen that $\text{var}(\hat{\delta}) > \text{var}(\bar{y}_T - \bar{y}_C)$, meaning that the variance on the estimated treatment effect using the regression adjustment in equation A.1 is actually larger than the variance of the simpler randomization-based estimate of the treatment effect, formed by simply contrasting the mean outcomes in the two groups. Unpublished simulations show that the variance penalty increases as the ratio of non-significant to significant covariates grows.⁹ There is a lack of good research on how large the ratio of cases to variables needs to be in order to guarantee that $\text{var}(\hat{\delta}) < \text{var}(\bar{y}_T - \bar{y}_C)$, but it appears that values of n/p less than 30 may be problematic. Eight of nine of the PACE sites have values of n/p in this potentially problematic range even after trimming the number of baseline predictors to 34 through the examination of their ability to explain measures derived from the National Student Clearinghouse about educational participation, persistence, and attainment (Fein, 2016b).

Based on this research, the team applied a slightly different approach to estimation for this report. The approach involved first estimating the influences of the baseline characteristics on the outcome under the control condition (equation A.2 below). The next step was to calculate

⁹ For example, with a sample size of 1000, when there are three covariates that explain 57 percent of the variation of the outcome and 97 covariates that are not relevant to prediction of the outcome, the standard error of the effect of treatment is 11 percent higher with OLS than with Koch’s method. (Austin Nichols, Abt Associates, unpublished simulations, 2016).

been expected under control conditions, as in equation A.3. These differences between actual and predicted outcomes are called “residuals.” The team then calculated the difference between average residual in the program group and the average residual in the control group, as in equation A.4. Equation A.5 gives the formula used to estimate standard errors on these impact estimates.

$$Y_i = X_i\beta + e_i, \tag{A.2}$$

$$\hat{r}_i = Y_i - X_i\hat{\beta}, \tag{A.3}$$

$$\hat{\delta} = \hat{\mu}_T - \hat{\mu}_C = \frac{\sum_i T_i \hat{r}_i}{\sum_i T_i} - \frac{\sum_i (1-T_i) \hat{r}_i}{\sum_i (1-T_i)}, \tag{A.4}$$

$$se(\hat{\delta}) = \sqrt{\frac{\sum_i T_i (\hat{r}_i - \hat{\mu}_T)^2}{\sum_i T_i - 1} + \frac{\sum_i (1-T_i) (\hat{r}_i - \hat{\mu}_C)^2}{\sum_i (1-T_i) - 1}}, \tag{A.5}$$

For survey-based outcomes subject to nonresponse, the team used a weighted version of this estimator (see Equation A.6).

$$\hat{\delta} = \frac{\sum_i w_i T_i \hat{r}_i}{\sum_i w_i T_i} - \frac{\sum_i w_i (1-T_i) \hat{r}_i}{\sum_i w_i (1-T_i)}, \tag{A.6}$$

where w_i is the nonresponse-adjustment weight for survey-reported outcomes.

This method is similar to the method developed by Koch, et al. (1998), who referred to it as nonparametric ANCOVA. Since then, most authors have referred to it as Koch’s estimator. The difference between Koch’s estimator and the method applied in this report is that Koch and co-authors fit equation A.2 on the entire sample rather than just the control sample. The main advantage of fitting A.2 just on the control sample is that the parameters are more easily interpretable when the null hypothesis is rejected. A secondary advantage is that, as Lesaffre and Senn (2003) demonstrated, Koch’s estimator can produce overly-liberal significance tests, meaning that the null hypothesis of no program effect is rejected too often. This occurs because the estimated standard errors on the estimated treatment effect using Koch’s method are too small. When the estimated standard errors are too small, random differences between the treatment and control groups are mistakenly classified as statistically significant evidence of program effects. Fitting A.2 on just the control sample will increase the estimated standard errors obtained in equation A.5 compared to what would be obtained by Koch’s estimator, but still smaller than what would be achieved with a pure randomization-based estimator.

Analysis confirmed that use of the modified Koch’s estimator slightly increased precision relative to both pure randomization and OLS (eq. A.1). The variance on the estimate of the impact of the program on the confirmatory outcome (average quarterly earnings in the sixth

Koch's estimator than it would have been with the OLS approach, and across a collection of confirmatory and secondary outcomes, the average variance reduction due to using the modified Koch's estimator instead of the OLS estimator was 2.2 percent.

Exhibit A.2 shows the regression coefficients from equation A.2 for the confirmatory outcome, average quarterly earnings for the sixth and seventh quarters after random assignment. Most of these covariates were selected based on a pooled analysis across all nine PACE sites of factors that predict various measures of success reported to the National Student Clearinghouse. Given the large sample size of Year Up PACE evaluation and the involvement of eight offices, additional covariates were added for office and cohort. Also, pre-randomization employment and earnings were added in for NDNH-defined outcomes only. Of 49 baseline covariates allowed into the model, twelve are predictive of average quarterly earnings at Quarters 6 and 7 for the control group sample. Specifically, being younger, female, black, planning at baseline to only attend school part-time are negatively associated with future earnings; while having one or more years of college but no degree or attending the Bay Area, Boston, Chicago, National Capital Region, or Puget Sound office is positively associated with future earnings. Additionally, every dollar of earnings during the quarter prior to random assignment is worth 39 cents of future earnings.

The team considered the alternative of OLS with a winnowed set of effectual covariates for each outcome at each PACE site but rejected doing so in favor of the greater transparency and convenience of using a common set of covariates for every outcome across the overall project.

Exhibit A.3 shows impacts on selected confirmatory and secondary outcomes before and after regression adjustment without weights.¹⁰ The impact estimates are all very similar. The standard errors with adjustment are modestly smaller with than without adjustment. For example, the standard error on the confirmatory outcome decreased by \$17, or 11 percent. This gain is equivalent to the expected improvement from increasing sample size by 27 percent. Reductions in standard errors for outcomes other than earnings were more modest. The reason is that no baseline variables predicted these other outcomes as well as pre-randomization earnings predicted future earnings.

¹⁰ See Exhibit B.3 in Appendix B for the impact of nonresponse-adjustment weights on the survey-based estimates.

Exhibit A-2: Coefficients for Baseline Characteristics as Predictors of Average Quarterly Earnings for Quarters 6 and 7 after Random Assignment: Year Up Control Group Members

Baseline Covariate	Estimate (\$)	Standard Error (\$)	p-Value
Intercept	1250.17	1699.42	0.462
Age			
20 or under	-2118.47	911.31	0.020
21 to 24	-1839.84	885.59	0.038
25 and older	0	na	na
Female	-385.98	226.71	0.089
Race/Ethnicity			
Hispanic, any race	-492.92	377.92	0.192
Black, non-Hispanic	-741.90	355.44	0.037
White, non-Hispanic	0	na	na
Another race, Non-Hispanic	-436.87	376.99	0.247
Living Arrangements			
Not living with spouse/partner or children	-367.53	632.85	0.562
Not living with spouse/partner, living with children	-83.18	666.62	0.901
Living with spouse/partner, not living with children	0	na	na
Living with spouse/partner and children	-6.41	408.27	0.987
Living with Parents	311.22	239.74	0.195
At Least One Parent with Some College	77.40	223.30	0.729
Usual High School Grades			
Mostly A's	-244.81	373.22	0.512
Mostly B's	-248.87	233.16	0.286
Mostly C's or below	0	na	na
Educational Attainment			
High school degree or less	0	na	na
Less than 1 year of college	-197.08	260.13	0.449
1 or more years of college	533.06	301.49	0.077
Associates degree or higher	317.65	539.90	0.556
Career Knowledge Index	-325.17	315.66	0.303
Psycho-Social Indices			
Academic Discipline Index	57.21	215.95	0.791
Training Commitment Index	144.95	223.23	0.516
Academic Self-Confidence Index	230.09	202.90	0.257
Emotional Stability Index	227.63	181.99	0.211
Stress Index	10.68	161.60	0.947
Public Assistance/Hardship Past 12 Months			
Received WIC or SNAP	6.64	234.38	0.977
Received public assistance or welfare	-193.76	448.96	0.666
Reported financial hardship	147.50	231.37	0.524
Current Work Hours			
0 to 19	0	na	na
20 to 34	-101.73	281.30	0.718
35 or more	243.26	383.20	0.526

(Continued)

Baseline Covariate	Estimate (\$)	Standard Error (\$)	p-Value
Expected Work Hours in Next Few Months			
0 to 19	0	na	na
20 to 34	-161.11	229.11	0.482
35 or more	-76.09	350.22	0.828
Life Challenges Index	-300.17	242.17	0.215
Plan to Attend School only Part-time if Admitted	-632.86	275.22	0.022
Employed during:			
Q4 before random assignment	-467.50	333.03	0.161
Q3 before random assignment	294.88	341.66	0.388
Q2 before random assignment	-520.54	371.07	0.161
Q1 before random assignment	-77.40	317.43	0.807
Earnings during:			
Q4 before random assignment	0.08	0.09	0.356
Q3 before random assignment	0.15	0.10	0.149
Q2 before random assignment	0.19	0.12	0.114
Q1 before random assignment	0.39	0.10	<0.001
Cohort:			
1 (January 2013 – March 2013)	361.07	719.12	0.616
2 (June 2013 – September 2013)	-517.41	446.17	0.247
3 (December 2013-March 2014)	140.93	305.81	0.645
4 (June 2014 – September 2014)	0	na	na
Office:			
A	1251.38	454.11	0.006
B	1944.03	409.01	0.000
C	1487.83	480.25	0.002
D	907.38	523.38	0.083
E	321.82	481.94	0.504
F	626.08	421.49	0.138
G	0	na	na
H	1572.99	541.33	0.004

SOURCE: Abt Associates calculations based on data from on data from the PACE Basic Information Form (BIF), and the PACE Self-Administered Questionnaire (SAQ).

NOTES: Model estimated with SAS/SURVEYREG procedure. Sample size=858. Family income inadvertently omitted from covariate list for impacts on NDNH outcomes but included as covariate for impacts on NSC and survey outcomes.

Exhibit A-3: Comparison of Selected Impact Estimates With and Without Adjustment for Baseline Imbalances

Outcome	Unadjusted Est (StdErr)	Adjusted Est (StdErr)
Confirmatory Outcome (NDNH)		
Average quarterly earnings for Quarters 6 and 7 after random assignment (\$)	1,857 *** (160)	1,895 *** (143)
Secondary Employment and Earnings Outcomes (NDNH)		
Employed as of a certain number of quarters after random assignment		
Q4 (%)	-6.88 (1.88)	-7.96 *** (1.79)
Q5 (%)	4.44 *** (1.78)	4.01 *** (1.64)
Q6 (%)	4.48 *** (1.72)	3.97 *** (1.67)
Q7 (%)	5.37 *** (1.80)	4.98 *** (1.71)
Average annual earnings during second year after random assignment Q4-Q7 (\$)	5,075 *** (547)	5,181 *** (474)
Sample Sizes (across treatment and control groups)		
2,544		
Other Secondary Outcomes (Survey, unadjusted for nonresponse)		
Indices of Self-Assessed Career Progress (average)		
Perceived career progress ^a	0.179 *** (0.033)	0.175 *** (0.032)
Confidence in career knowledge ^b	0.104 *** (0.030)	0.098 *** (0.028)
Access to career supports ^c	0.098 *** (0.014)	0.095 *** (0.014)
Indicators of Career Pathways Employment (%)		
Working in a job paying \$15/hour or more ^d	32.19 *** (2.00)	31.46 *** (1.91)
Working in a job requiring at least mid-level skills	28.32 *** (2.00)	28.11 *** (1.95)
Sample Sizes (across treatment and control groups)		
1,939		
NSC-Reported Educational Progress		
Any enrollment during Quarters 0 through 3 (%)	32.68 *** (1.97)	31.21 *** (1.84)
Any enrollment during Quarters 4 through 7 (%)	-0.586 (0.193)	-7.17 (1.82)
Cumulative FTE months of enrollment by Q7	1.441 *** (0.203)	1.272 *** (0.191)
Sample Sizes (across treatment and control groups)		
2,496		

SOURCE: Abt Associates calculations based on data from PACE short-term follow-up survey.

NOTES: Standard errors on estimated impacts are shown in parentheses. Adjusted impact estimates and associated standard errors were prepared with the modified Koch's estimator, as defined equations (A.4) and (A.5). Statistical significance levels, based on one-tailed t-tests tests of differences between research groups, are summarized as follows: *** statistically significant at the one percent level; ** at the five percent level; * at the ten percent level.

^a Three-item scale tapping self-assessed career progress; response categories range from 1=strongly disagree to 4=strongly agree.

^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.

^d Year Up's wage performance standard.

Appendix B: Unemployment Insurance Wage Data

A large collection of social program evaluations have relied on earnings data resulting from state unemployment insurance (UI) tax filings by employers. 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 the NDNH’s virtues are that, unlike state data, it captures earnings for research participants who move to another state during the follow-up period.

The federal Office of Child Support Enforcement (OCSE) in the Department of Health and Human Services’ Administration for Children and Families operates the NDNH. The NDNH contains new hire, quarterly wage, and unemployment insurance information submitted by State Directories of New Hires, employers, and State Workforce Agencies. The 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 concealed tips, self-employment, firms’ employment of independent contractors, and informal employment.¹¹

B.1 Data Collection Process

The primary purpose of the NDNH is to assist state child support agencies locate noncustodial parents, putative fathers, and custodial parents to establish paternity and child support obligations, as well as to enforce and modify orders for child support, custody, and visitation. It is also used by state unemployment insurance agencies and the federal Social Security Administration (SSA) to identify overpayments of benefits. However, subject to federal law, regulation, guidance, and other requirements to protect data privacy and security,¹² 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.

OPRE and OCSE negotiated a memorandum of understanding (MOU) allowing access to NDNH data for the PACE evaluation (including Year Up). Among other provisions, the MOU dictates what self-reported data from study subjects can be merged with NDNH data, the computing environment where these merges are conducted, and procedures for review of tables prior to

¹¹ According to the 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>.

¹² 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.

release. As a contractor to OPRE, the evaluation team at Abt Associates must comply with the terms of this MOU.

The evaluation team transmits match request files to OCSE on a quarterly basis. These match request files contain the names and social security numbers of PACE research participants. OCSE verifies with the SSA that the reported SSNs belong to the named persons. For those SSNs that pass this test, the OCSE copies NDNH records for the quarter and the preceding seven quarters to a secure folder on the ACF server.¹³ (Ordinarily, these records would be destroyed after two years.) These copied records contain a pseudo-SSN. They are stripped of all personal identifiers. States are required to submit earnings records within 4 months, but there are stragglers and corrections. The evaluation team required six elapsed months in order to assume that quarterly earnings were reasonably completely reported to NDNH.

Once the PACE evaluation team is ready to analyze the collected data, it submits a “pass-through” 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 of this report). OCSE then strips the personal identifiers out of the pass-through file and replaces the actual SSN with the same pseudo-SSN they had previously assigned to the archived wage records. This pseudo-SSN allows the PACE evaluation team to merge program and self-reported data with the quarterly wage data in order to estimate program impacts on earnings and employment.

B.2 Data and Measures

Abt Associates received 21 files for use in this report, one for each quarter between Q1 2013 and Q1 2018. Random assignment of the first cohort of Year Up applicants started in January of 2013 and random assignment of the fourth (and final) cohort ended in September of 2014. Given the six-month lag in processing of employer reports by the states and transfer of state data to OCSE, wage records were available through Q3 2017; this means that the evaluation team had 18 post-randomization quarters of earnings data for the first cohort and 12 post-randomization quarters of earnings data for the final cohort. In addition, the evaluation team had eight quarters of pre-randomization data for the entire sample (analyses included only the four most recent quarters in regression-adjustment models).

Of the 2,544 people randomized as part of the PACE Year Up evaluation, 2,496 reported names and social security numbers that OCSE deemed to be of sufficient quality for their matching purposes. Analyses in this report thus are based on the 98.1 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, the team defined sample members as not working when there was no match to wage records in a given quarter.

Each quarter, the evaluation team submitted a match request file to OCSE that contained the names and social security numbers for everyone randomized to that date. For those where the SSNs and names were aligned, OCSE returned earnings data for the eight most recent quarters

¹³ Those participants who are not matched in the SSA database are considered “missing” for these purposes, because their employment records are not available.

in the NDNH, which is lagged by two quarters from the date of the match. This meant that the evaluation team had up to eight wage reports for each quarter. The team used the last version for each quarter within a window. For example, for earnings in the second quarter of 2014, the team 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, the team assumed that these reports reflected either distinct payments by the same employer or payments from different employers. Consistent with the logic discussed in Appendix F, the team reviewed quarterly earnings for any values that were clearly impossible, but failing to find any such values, did not discard or top-code any large earnings amounts.

The team calculated two outcomes for each quarter: a binary indicator of any earnings and the total reported wages for the quarter. The result was two series of 14 outcomes for each person (employment and earnings for Quarters -2 to +11. In addition, the team formed a quarterly average for Quarters 6 and 7 after random assignment (the confirmatory outcome, established to align with the Year Up logic model) and annual averages for Quarters 0-3, 4-7, and 8-11.

Appendix C: National Student Clearinghouse Data

The National Student Clearinghouse (NSC) is a national database of college enrollment records designed to aid the administration of student loans programs but can be a useful tool for education researchers. A few caveats need to be noted, however. First, given the focus on loan administration, it does not cover schools that are not Title IV schools, the set of schools approved for federal student loans by the US Department of Education. Second, information on outcomes other than enrollment tends to be less reliable. Notably, standards and practices governing credential reporting are inconsistent across schools.

C.1 Coverage

Exhibit C-1 shows the percentage of schools providing records to the NSC by year and by type of school. As shown, coverage of public two-year and four-year schools was over 95 percent. Coverage was lower among private not-for-profit four-year schools, considerably lower among private for-profit four-year schools, and very low for private two-year schools (both for-profit and not-for profit). Since it is plausible that people randomly denied access to Year Up might be more likely to attend for-profit schools, there is some possibility for upward bias in estimated college enrollment impacts based on NSC data.

Exhibit C-1: NSC School-level Cooperation Rates by School Control and Level from 2012 through 2016

Type and Control of School	Year			
	2013 (%)	2014 (%)	2015 (%)	2016 (%)
Public, 4-year	99.2	99.4	99.5	99.6
Private, not-for-profit, 4-year	93.6	95.2	95.8	96.1
Private, for-profit, 4-year	74.4	79.9	81.7	81.0
Public, 2-year	99.1	99.2	99.4	99.5
Private, not-for-profit, 2-year	39.5	40.8	40.4	42.1
Private, for-profit, 2-year	19.7	28.1	26.7	26.6

SOURCE: National Student Clearinghouse https://nscsearchcenter.org/wp-content/uploads/NSC_COVERAGE.xlsx.

Analyses of NSC data in this report are limited to enrollment records obtained from 2000 forward. The PACE evaluation team obtained informed consent from all study participants to have their records from NSC released to the evaluation team. The team negotiated a contract with NSC to match relevant NSC records to the study subjects using social security numbers and names. The abstracted records were then sent by encrypted secure methods to the evaluation team who have used them under tight security conditions.

C.2 Data and Measures

Counting the quarter during which random assignment occurred as Quarter 0, the evaluation team obtained an abstract from NSC in October of 2016 covering enrollment through Quarter 9 for all 2,539 study subjects (1,668 in the treatment arm and 871 in the control arm of the study) and through Quarter 10 for the 2,004 study subjects (1,321 T and 683 C) enrolled by March 31, 2014.

Records arrived in a spell format with starting and ending dates. The evaluation team translated these first into a set of person-month level records, reconciling multiple and conflicting spells as seemed most sensible. The team derived two variables for each person month. The first was a simple binary indicator of “any enrollment.” The second was a measure of full-time equivalent (FTE) enrollment that took the values 1 (for full-time enrollment), 0.75 for three-quarter-time enrollment, 0.5 for half-time enrollment, 0.25 for some but less than half-time enrollment, and 0 for no enrollment.¹⁴ To translate these to person-quarter-level outcomes, a student was counted as enrolled for the quarter if they were enrolled at all in any of the three months, and FTE enrollment was calculated by summing their total FTE months for the quarter.

¹⁴ Because informed consent had been collected from all study participants, NSC shared full/part-time status for everyone in the sample, something that is not otherwise shared with researchers.

Appendix D: 18-Month Follow-up Survey Data

Abt's survey group conducted the follow-up survey approximately 18 months after random assignment in all nine PACE sites, including Year Up. Starting in the 15th month after random assignment, interviewers initiated attempts to contact sample members. The target population for the survey included the full sample of young adults randomly assigned in PACE. Interviews lasted about an hour and were conducted by trained interviewers using computer-assisted technology. Initial contact attempts were by phone from a centralized facility. In the second phase, local interviewers made home visits to try to interview initial non-respondents. On average, interviews occurred 17.5 months after random assignment. The average follow-up duration was very similar for the treatment and control groups (less than two weeks longer for the latter than for the former).

The survey collected measures for a variety of outcomes in domains represented in PACE's general theory of change. Among these outcomes were different aspects of training experiences, self-assessed career progress, psycho-social skills, career networks, employment and earnings, financial hardship and other stresses, dependence on public anti-poverty programs, and family formation and childbearing. Section D.1 identifies the outcomes analyzed in this report. A full facsimile of the questionnaire is available at the OMB website.¹⁵ Section D.2 discussed imputation procedures for missing items, and Section D.3 discusses unit nonresponse and a weighting adjustment that was made for it.

D.1 Outcome Measures

Exhibit D.1 provides details on specifications for outcomes representing program experiences, analyzed in Chapter 5. Exhibit D.2 provides similar information for the outcomes analyzed in Chapter 6.

D.2 Imputation for Item Nonresponse

This section documents the evaluation team's response to two sources of missing data affecting survey outcomes. First, initial data quality assessment revealed that a small fraction of respondents who initially indicated receiving some education and training since random assignment (in A1 and A1a) did not answer a series of subsequent filter questions (A10) on types of classes took (ESL, adult basic education, classes for college credit, noncredit occupational training, life skills classes). Second, all outcomes were affected by at least some missing data where respondents either declined to answer a question or gave an answer of "don't know."

Concerning the first issue, checks against Year Up administrative records for members of the treatment sample confirmed that the majority of those with A1=yes (some training since random assignment) but A10=no for every type of classes did, in fact, receive (and complete) training at Year Up. For Year Up, the discrepancy affected fewer than 4 percent of respondents and occurred at similar rates for treatment (five percent) and control (three percent) group members.

¹⁵ <http://www.reginfo.gov/public/do/DownloadDocument?objectID=38944602>.

Exhibit D-1: Details on Specifications for Survey-Based Outcomes in Chapter 5

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Experiences of Everyone Randomized		
Received education or training since random assignment	Two question format with slightly different wordings to try to get all training spells reported	A1, A1a
Cited financial support as a challenge in school or persistence	Reported money troubles as reason for not continuing studies, not currently studying, or never starting studies; or reported that it was very or somewhat difficult to obtain adequate financial support to continue their studies	A11a, A14a, A23a, A26a, A35, A59, A60
Ever receiving help with:	This was asked of everyone, even those with no training since random assignment.	A62
Arranging supports to meet school, work or family responsibilities		
Career counseling		
Job search or placement		
Experiences of Everyone Receiving Any Education or Training		
Classes at First Place Spent Most/All Time on:		
Listening to lectures		A47a
Group discussion		A47b
Group projects		A47c
Individual projects		A47d
Indicate that classes at first place often used active learning methods	Responses to three positively worded items from 6-item battery were reverse scaled (1=none of the time, 4=all the time) and then averaged. Three negatively worded items were not used because they did not exhibit the expected negative correlations with the positively worded items. Anyone with an average of 2.5 or larger was counted.	A47b-A47d
Strongly agree that classes at first place were relevant to life/career	Strongly agrees that, "These classes were relevant to my career interests," or strongly disagrees that, "These classes did not relate to much of anything else in my life."	A46c, A46d
At first place, received:		
Career counseling		A36d
Academic advising		A36a
Financial aid advising		A36b
Tutoring		A36c
Help arranging supports for school or work		A36f
Grants/scholarships	A Pell grant or other government grant or scholarship—not counting loans you have to pay back, Must indicate in A31 that funds were used with for tuition, other school related expense, or living expenses.	A30g, A31
Loan(s)	Loans in your own name or loans in your parents' names. Must indicate in A31 that funds were used with for tuition, other school related expense, or living expenses.	A30e, A30f

(Continued)

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
At first place,:		
Perceived a great deal of emphasis on community	People who responded "a great deal" were counted.	A37
Offered opportunities for related work experience	Question was asked about each place attended since random assignment, but only information on first place was analyzed. Enrollment dates were used to determine first place attended since random assignment.	A38
Received job search or placement assistance		A36e
Life Skills Instruction at Any Place Among Those Receiving Any E & T		
Received life skills instruction since random assignment		A10e
Experiences of Everyone Receiving Any Life Skills Instruction at Any Place		
Reported a Great Deal of Attention in These Classes Paid to:		
<i>Study Skills</i>		A53b
<i>Critical Thinking</i>		A53d
<i>Time Management</i>		A53g
<i>Managing Stress/Anger</i>		A53j
<i>Staying Motivated</i>		A53k
<i>Acting Professionally</i>		A53l
<i>Communicating Well</i>		A53i
<i>Working in Groups</i>		A53h
<i>Managing Finances</i>		A53m
<i>Career Planning</i>		A53a
<i>Job Search</i>		A53c
<i>Dealing with Other Life Problems</i>		A53n

Exhibit D-2: Details on Specifications for Survey-Based Outcomes in Chapter 6

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Exploratory Indicators of Earnings and Employment		
Current employment		E1
Quarterly earnings	Multiplied hourly wages by hours worked by 13 if hourly wage was reported. Scaled up earnings for periods other than hours appropriately.	E1a, E2
Hours worked last week		E1a
Hourly wages if employed		E2
Indicators of Career Pathways Employment		
Working and \$15 or more	Threshold of \$15/hour corresponded to Year Up goals.	E2
Working 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 one of the Department of Labor Standard Occupational Classification (SOC) codes. The team then looked up the Job Zone ¹⁶ for each SOC code in the BLS O*NET system. ¹⁷ 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 of 3—occupations that need an intermediate level preparation—seemed a reasonable goal for graduates of Year Up. This Job Zone 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.”	E3, E4, E5
Working in a Year Up target occupation	Three open-ended questions about the kind of work done, the usual activities completed, and the job title were coded into one of the Department of Labor Standard Occupational Classification (SOC) codes. SOC codes were then classified on the basis of whether or not they corresponded to Year Up target occupations. To make this determination, PACE analysts compared survey-based SOC codes for treatment group members with Year Up staff members’ determinations of whether each sample member was “working in target occupation,” as recorded in the program’s administrative database (Salesforce®). These staff determinations were available for program completers at the four-month post-graduation mark.	E3, E4, E5

(Continued)

¹⁶ <https://www.onetonline.org/help/online/zones> [last accessed September 12, 2016]

¹⁷ <https://www.onetonline.org/> [last accessed September 12, 2016]

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Indices of Self-Assessed Career Development		
Perceived career progress	<p>This was a new scale created for PACE. It is a 3-item scale of self-assessed career progress; response categories range from 1='strongly disagree' to 4='strongly agree'. It was designed specifically to measure an individual's sense of progress 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 	C5, C6
Confidence in career knowledge	Same as at baseline, as described in Exhibit A-1.	C3
Access to career network	<p>This was a new scale created for PACE. It is a 6-item yes/no scale, counting number of types of career-supportive relationships in workforce and education settings. The motivation for creating this scale was the theory richer social networks are one of the benefits of higher education (e.g., Goldrick-Rab and Sorenson, 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? 	C2
Other College Outcomes		
Number credits received	Summed across schools verified in IPEDS to be degree-granting institutions	A4, A5, A25a
Received a credential from a college	The survey had separate questions about credentials awarded for regular for-credit classes and for noncredit occupational classes. If the respondent indicated receiving either type of credential, then this variable was coded as 1 (for yes); otherwise, it was coded as 0 (for no). The survey did not ask for credentials awarded as a result of ESL, ABE, or life-skills classes. The survey also reported school names. Only credentials awarded by schools listed as degree-granting by IPEDS were counted for this outcome.	A4, A5, A22, A23, A27e, A27f
Received a credential from another education-training institution	Same as at colleges.	A22, A23, A27e, A27f
Received a credential from a licensing/certification body	The survey asked about the highest level of occupation training completed. One of the possible answers was "a professional, state or industry certification, license or credential." If the respondent picked this level, then there was a follow-up question about the year of award. If the year of award was the same as the year of random assignment or later, then the person was coded as having earned such a credential.	A56
Received a credential from any source	See cells above for receipt of credentials from colleges, for other education training institutions, and from licensing/certification bodies. If a student had obtained any of these, he or she was classified as having received a credential.	A22, A23, A27e, A27f, A56

(Continued)

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Indicators of Financial Status and Related Strains		
Has health insurance	Through work, Medicaid, or other	E8
Household receiving cash/in-kind supports	A yes response on one or more of 11 listed federal and state anti-poverty programs <ul style="list-style-type: none"> • Temporary Assistance to Needy Families (TANF) • Supplemental Nutrition Assistance Program (SNAP) • Social Security Income (SSI) • Unemployment Insurance (UI) • Medicaid • Subsidized child care • Section 8 / Public housing • Low Income Home Energy Assistance Program (LIHEAP) • Free or reduced lunch program • General assistance (using local state name) • Women, Infants and Children (WIC) program 	E7
Experienced financial hardship in the last 12 months	Same as at baseline. See Exhibit A-1.	D1, D2
Index of life challenges	Slight modification to scale with same name described in Exhibit A-1. Version at baseline listed 6 situations that could interfere with school, work, job search, or family members. Version at follow-up included a seventh situation: <ul style="list-style-type: none"> • Spending time with friends 	D3
Index of perceived stress	Existing scale from Cohen et al. (1983). This scale was used for a second time in the follow-up survey. It was used initially in in the BIF. The 4-item scale captured perceived stress. The response categories ranged from 1='never' to 4='very often'.	D4
Indicators of Psycho-Social Skills and Family Formation		
Psycho-Social Skills		
Grit	Existing scale from Duckworth, et al. (2007). The 8-item scale captures persistence and determination. Response categories ranged from 1 (strongly disagree) to 4 (strongly agree).	B3
Savvy	Existing scale from Le, et al. (2005). The 13- -item scale includes response categories that range from 1='strongly disagree' to 6='strongly agree'.	B5
Core self-evaluation	Existing scale from Judge (2009). The 12- item scale response categories ranged from 1='strongly disagree' to 4='strongly agree'.	B6
Family Structure		
Living with spouse	Unmarried partners not counted	F1a
Had child since random assignment/ currently pregnant	Not asked of men not living with spouse or partner. Analyses therefore restricted to women.	F2, F3
Not living with parents		F1e

To adjust for these missing data, the team imputed new responses for A10 using a nearest neighbor hot deck procedure (Andridge & Little, 2010).¹⁸ The hot deck involves “binning” and sorting. Within a bin, the procedure matches each case that is missing an outcome to the nearest complete case with respect to the sort. This hot deck imputation procedure matched spells with consistent responses to A10 (*consistent spells*) to spells with inconsistent responses to A10 (*inconsistent spells*). The team used site and treatment status to define the bins and the modeled propensity of a spell being consistent to define the sorting variable. To model the propensity that a spell would be consistent, the team searched a large potential set of predictor variables from baseline variables and from sections of the follow-up survey for which A10 was not a filter question. The team included interactions as well as main effects. The team conducted this search and fit the final model on a pooled dataset including observations from Year Up, as well as five other PACE programs to boost power.¹⁹ The final imputation model used 24 variables and interactions from the survey.

In the course of imputing A10, the team kept track of the ID of the consistent spell that was matched to each inconsistent spell. After imputation of A10 was complete, the team then filled in responses to the detailed questions (A11-A29) filtered by A10 by copying the responses for the consistent spell that had been matched to the inconsistent spell.

In response to the section issue—the common problem of small fractions missing on most questions due to refusals and don’t knows—the team for the most part simply omitted people with such responses from the relevant analyses. This was done separately for each outcome, meaning that the maximum number of usable responses was used for estimating the impact of each outcome. However, for training hours the team imputed responses for each type of classes at each school the respondent attended. This imputation allowed the team to sum training hours across schools and types of classes without having high missing data rates on the sums because of scattered item missingness. To carry out this imputation, the team used SUDAAN/IMPUTE, as discussed in Section A.1 for missingness of baseline covariates. This random matching was constrained to occur within strata defined by treatment status, site, type of training, and self-reported completion status of the spell.

D.3 Nonresponse Analysis and Weighting

The 18-month follow-up survey obtained a higher response rate in the treatment group (78 percent) than in the control group (73 percent). This section describes analyses of, and weights developed to adjust for, nonresponse.

Exhibit D.3 compares distributions on baseline characteristics for all sample members and survey respondents. There was only one significant imbalance (using a threshold of 0.10 for statistical significance) on the full sample. The same statistic (at least one parent went to college) also was imbalanced for the (unweighted) survey respondent sample.

¹⁸ If A10e was answered “no” or was not answered, then items A49-A51 were skipped. The team decided not to impute values for these items in the cases where A10e was imputed to have a value of “yes,” as A49-A51 do not provide important outcomes for PACE impact analyses.

¹⁹ Data collection was completed at three sites substantially earlier than the rest, and so they were processed together. The final six sites were then processed together for purposes of A10 imputation.

The upper panel of Exhibit D.4 compares regression-adjusted impacts on college outcomes from NSC records for the full and respondent samples. Standard errors on the unweighted respondents are larger because of the reduced sample sizes, but point estimates for impacts are very similar. For the three outcomes in the table where Year Up is estimated to have had significant impacts, the point estimates are slightly larger on the unweighted respondent sample than on the full sample, perhaps reflecting a greater willingness of people with favorable outcomes to respond to the survey request.

In response, the team developed and applied weights to adjust for nonresponse, based on statistical models of the association between baseline characteristics and response probabilities within each of the two randomly assigned groups. Covariates also included several measures of college enrollment and credential receipt over the follow-up period. These methods are common in survey research.

The main steps in constructing weights included:

1. Winnow the list of potential covariates that are statistically significant in a logistic regression model for response status.²⁰ Do this separately for treatment and control cases. This approach identified age, family structure, and living with parents as significant predictors of response status in the treatment sample. The set of significant predictors in the control sample consisted of family structure, typical high school grades, and stress.
2. Using the winnowed list of potential covariates, estimate the response propensity for each member of the treatment and control sample—both for respondents and non-respondents.
3. Sort the sample in each study arm by the estimated response propensity, and then divide the sample into five equal-size groups (quintiles).
4. Within each arm and quintile, calculate the empirical response rate. Invert it to calculate the nonresponse-adjusted weight.

The last column in Exhibit D.3 shows that the weighting had little effect on baseline imbalances. With nonresponse weighting, there are only two statistically significant imbalances. The last column in the lower panel of Exhibit D.4 shows that weights had little effect on impacts on college enrollment outcomes recorded in the NSC. To guard against possible biases in survey measured outcomes, the team nonetheless applied the nonresponse-adjusted weights in all analyses of survey-based outcomes for this report.

²⁰ The team used the stepwise search option in SAS/LOGISTIC for this purpose with a p-value to enter the model of 0.20 and a p-value to stay in the model of 0.10.

Exhibit D-3: Baseline Balance on Full Sample, Unweighted Respondent Sample, and Weighted Respondent Sample

	Year Up Baseline Characteristics								
	All Participants			Survey Respondents, Unweighted			Survey Respondents, Weighted		
	Treatment	Control	p-value	Treatment	Control	p-value	Treatment	Control	p-value
Age (%)	-	-	0.811	-	-	0.882	-	-	0.835
20 or under	43.2	42.1	-	43.5	44.5	-	43.4	42.4	-
21 to 24	56.0	57.3	-	55.6	54.7	-	55.7	56.9	-
25 or older	0.8	0.7	-	0.9	0.8	-	0.9	0.7	-
Gender (%)	-	-	0.992	-	-	0.991	-	-	0.799
Female	41.0	41.0	-	41.4	41.4	-	41.2	40.6	-
Male	59.0	59.0	-	58.7	58.6	-	58.8	59.4	-
Race/Ethnicity	-	-	0.557	-	-	0.728	-	-	0.750
Hispanic, any race	31.6	31.1	-	31.0	31.3	-	31.2	31.0	-
Black, non-Hispanic	53.9	54.2	-	54.3	55.6	-	54.3	55.8	-
White, non-Hispanic	7.7	10.1	-	7.6	9.4	-	7.7	9.5	-
Another race, non-Hispanic	11.6	10.9	-	12.1	10.4	-	11.8	10.3	-
Living Arrangements (%)	-	-	0.452	-	-	0.069	-	-	0.192
Not living with spouse/partner or children	87.1	85.8	-	87.5	84.5	-	87.6	85.9	-
Not living with spouse/partner, living with children	6.6	6.2	-	6.7	6.5	-	6.6	5.9	-
Living with spouse/partner, not living with children	4.2	5.1	-	3.8	5.5	-	3.8	5.0	-
Living with spouse/partner and children	2.1	2.9	-	2.0	3.5	-	2.0	3.2	-
Living with Parents (%)	68.8	67.7	0.564	68.4	68.4	0.976	68.4	68.7	0.895
<i>At Least One Parent with Some College (%)</i>	<i>54.0</i>	<i>58.0</i>	<i>0.064</i>	<i>55.5</i>	<i>60.3</i>	<i>0.056</i>	<i>54.8</i>	<i>60.4</i>	<i>0.027</i>
Usual High School Grades (%)	-	-	0.249	-	-	0.599	-	-	0.655
Mostly A's	10.0	11.1	-	10.4	11.8	-	10.3	11.6	-
Mostly B's	48.6	50.8	-	49.4	49.8	-	49.5	49.5	-
Mostly C's or Below	41.4	38.1	-	40.2	38.5	-	40.3	38.9	-
Educational Attainment (%)	-	-	0.791	-	-	0.551	-	-	0.645
Less than a high school degree	0.7	0.5	-	0.7	0.2	-	0.7	0.2	-
High school degree or equivalent	51.0	53.3	-	50.7	51.9	-	51.2	51.5	-
Less Than 1 year of college	22.3	21.5	-	21.7	22.3	-	21.7	22.4	-
1 or more years of college	22.9	21.8	-	23.8	22.3	-	23.3	22.5	-
Associates degree or higher	3.1	2.9	-	3.1	3.3	-	3.1	3.4	-

Continued

Year Up Baseline Characteristics									
	All Participants			Survey Respondents, Unweighted			Survey Respondents, Weighted		
	Treatment	Control	p-value	Treatment	Control	p-value	Treatment	Control	p-value
Received Vocational or Technical Certificate or Diploma (%)	18.9	17.3	0.349	19.4	18.1	0.521	19.3	18.2	0.551
Career Knowledge Index (average of items)	0.46	0.47	0.625	0.45	0.46	0.432	0.45	0.46	0.414
Psycho-Social Indices									
Academic Discipline Index	5.28	5.27	0.671	5.26	5.24	0.414	5.27	5.23	0.324
Training Commitment Index	5.52	5.50	0.232	5.52	5.49	0.263	5.52	5.48	0.214
Academic Self-Confidence Index	5.04	5.07	0.226	5.03	5.04	0.673	5.03	5.04	0.667
Emotional Stability Index	5.33	5.32	0.987	5.31	5.31	0.990	5.32	5.32	0.970
Social Support Index	3.34	3.36	0.135	3.33	3.35	0.258	3.33	3.35	0.239
Stress Index	2.21	2.18	0.315	2.22	2.19	0.404	2.23	2.20	0.408
<i>Depression Index</i>	<i>1.60</i>	<i>1.57</i>	<i>0.103</i>	<i>1.61</i>	<i>1.58</i>	<i>0.120</i>	<i>1.61</i>	<i>1.57</i>	<i>0.093</i>
Income (%)	–	–	0.533	–	–	0.940	–	–	0.913
Less than \$15,000	37.3	36.7	–	37.2	37.3	–	37.3	36.8	–
\$15,000-\$29,999	25.0	27.1	–	25.5	26.1	–	25.5	26.4	–
\$30,000 or More	37.7	36.2	–	37.3	36.6	–	37.2	36.7	–
<i>Mean (\$)</i>	<i>27,287</i>	<i>26,528</i>	<i>0.443</i>	<i>27,245</i>	<i>26,986</i>	<i>0.836</i>	<i>27,165</i>	<i>27,105</i>	<i>0.961</i>
Public Assistance/Hardship Past 12 Months (%)									
Received WIC or SNAP	32.6	33.1	0.756	32.9	33.8	0.689	33.1	33.4	0.913
Received public assistance or welfare	6.3	7.3	0.352	6.5	7.0	0.722	6.6	6.9	0.791
Reported financial hardship	29.4	30.3	0.578	30.0	30.3	0.884	30.1	30.5	0.859
Current Work Hours (%)									
0	47.9	47.1	0.490	48.5	48.1	0.352	48.6	47.6	0.359
1 to 19	10.3	11.0	–	10.9	11.6	–	10.9	11.6	–
20 to 34	27.4	25.3	–	26.9	24.1	–	26.9	24.4	–
35 or more	14.5	16.6	–	13.6	16.2	–	13.6	16.3	–
Expected Work Hours in Next Few Months (%)									
0	35.9	37.1	0.866	35.6	37.1	0.722	35.4	36.7	0.774
1 to 19	23.1	22.7	–	24.3	22.0	–	24.3	22.2	–
20 to 34	31.5	30.3	–	30.9	31.1	–	31.1	31.2	–
35 or more	9.4	9.9	–	9.3	9.9	–	9.3	9.9	–

Continued

Year Up Baseline Characteristics									
	All Participants			Survey Respondents, Unweighted			Survey Respondents, Weighted		
	Treatment	Control	p-value	Treatment	Control	p-value	Treatment	Control	p-value
Life Challenges Index (mean response on 1-5 scale)	1.47	1.45	0.264	1.47	1.45	0.468	1.47	1.45	0.406
Owns a Car (%)	28.7	28.9	0.959	28.9	28.2	0.752	28.7	28.2	0.836
Has both Computer and Internet at Home (%)	84.1	86.5	0.111	84.4	86.9	0.145	84.2	87.0	0.103
Ever arrested (%)	16.6	15.5	0.502	16.4	14.7	0.336	16.5	15.2	0.471
Sample Sizes	1,669	875	–	1,301	638	–	1,301	638	–

SOURCE: Abt Associates calculations based on data from the PACE Basic Information Form (BIF), the PACE Self-Administered Questionnaire (SAQ), and response status to the PACE short-term follow-up survey.

NOTES: SAS/SURVEYFREQ used to test for significant imbalances for categorical variables. SAS/TTEST was used to significant imbalances for other variables.

Exhibit D-4: Comparison of Selected Impact Estimates for the Unweighted and Weighted Survey Samples

Outcome	Full Sample	Survey Respondents	
		Unweighted Estimate	Weighted Estimate
Average Quarterly Earnings in Quarters 6 and 7 (Standard Error)	\$1,895 *** (\$143)	\$2,065 *** (\$162)	NA
College Outcomes			
Percent with any enrollment in Quarters 0-3 (Standard Error)	31.21 *** (1.84)	31.61 *** (2.12)	31.46 *** (2.10)
Percent with any enrollment in Quarters 4-7 (Standard Error)	-7.17 (1.82)	-7.03 (2.11)	-6.77 (2.08)
Cumulative FTE months of enrollment by Q7 (Standard Error)	1.27 *** (.191)	1.28 *** (.225)	1.28 *** (.215)
Sample Size	2,539	1,939	1,939

SOURCE: Outcomes data are based on quarterly wage records obtained in a match to the National Directory of New Hires and on college records in the National Student Clearinghouse.

NOTES: Statistically significant in a one-tailed test: * at the 90% level, ** at the 95% level, *** at the 99% level.

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

This evaluation compares estimates of employment and earnings impacts based on UI records and survey self-reports. Barnow and Greenberg (2015) review findings from evaluations including both data sources. Although average survey-reported earnings tend to be higher than average total UI earnings, impact estimates still may be consistent (Kornfeld and Bloom 1999).

The top panel in Exhibit E-1 shows that impact estimates for Year Up derived from the two sources (presented earlier in Chapter 6, Exhibit 6-1) agree closely. The estimated UI records-based impact of \$1,895 impact for average earnings in Quarters 6 and 7 (the confirmatory outcome) is very similar to the \$1,970 impact estimate based on 18-month follow-up survey data.²¹

The small difference between impact estimates from NDNH and the survey (\$75) may result in part from underreporting of wage records to the federal NDNH database by Washington State. The Abt team identified a number of calendar quarters with lower NDNH record counts for Washington State than expected based on counts in adjacent quarters.²²

To confirm this suspicion, we used survey data for Year Up's Seattle office to adjust the corresponding NDNH estimate. The adjustment applied a two-step procedure to office-level statistics. The first step was to run an ordinary least squares regression on point estimates for impacts from the seven non-Washington offices. Specifically, analysts regressed NDNH-based impacts on the corresponding survey-based estimates for each office ($R^2 = 89\%$). The second step involved multiplying the survey-based impact estimate for Washington by the regression coefficient relating survey to NDNH estimates in the remaining offices. The result was an adjusted NDNH estimate for Washington. The same procedure was used to estimate mean NDNH earnings for the Washington control group.

The adjusted overall NDNH impact was \$1,990 (not shown in exhibit)—nearly identical to the survey estimate (\$1,970)—suggesting that the under-reporting of Washington data explains most of the NDNH-survey discrepancy. As another check on the NDNH-based estimate, we re-ran impacts for the confirmatory earnings outcome excluding sample members from Year Up's Seattle site (10 percent of the sample). The resulting impact estimate (\$1,932) was very close to the adjusted full-sample (\$1,990) and survey (\$1,970) estimates.

The second panel of Exhibit E-1 shows that NDNH-based employment estimates are slightly higher than survey-based estimates for treatment group members (80 and 74 percent, respectively) but nearly identical for control group members (76 and 74 percent), leading to somewhat different estimated employment impacts. The differences between NDNH and survey-based measures could reflect sampling error, or they might result from differences in

²¹ The survey figures convert the available survey measure—earnings in the prior week (calculated as hourly wage times number of hours worked)—to a calendar quarter-level estimate by multiplying by 13 (the average number of weeks in a quarter).

²² ACF, including the federal Office of Child Support Enforcement (which maintains the NDNH), is working with the State to obtain improved data for future PACE analyses.

time periods covered by each measure. The NDNH-based measure captures quarterly earnings, while the survey-based estimate uses a single week’s worth of earnings to estimate quarterly earnings (e.g., multiplying by 13, the average number of weeks in a quarter). NDNH thus reflects the cumulative effect of brief jobs to a greater extent than the (weekly-based) survey figure. If treatment group members were more likely to have very short job spells after graduation than control group members had in the same follow-up period, we would expect to see higher fractions of the former employed in NDNH than in the survey but smaller differences in average earnings (because the additional jobs were very brief). Higher job turnover is quite plausible given the strong drive to help participants find jobs following Year Up graduation and anecdotal reports that multiple offers and switches were not uncommon. Supporting this interpretation, employment impacts fade while substantial earnings impacts persist in the third follow-up year (see graphs in Chapter 6, Section 6.1).

Exhibit E-1: Impacts on Earnings and Employment around Follow-up Quarters 6-7 Based on Wage Records and Self-Reports

Outcome	Treatment	Control	Impact
Quarterly Earnings			
Average NDNH earnings in Quarters 6-7 (\$)	5,454	3,559	1,895 ***
(Standard error)	–	–	(143)
Self-reported earnings (\$)	5,907	3,937	1,970 ***
(Standard error)	–	–	(183)
Employment			
Average percent with employer-reported wages in Quarters 6-7	80.3	75.8	4.5 ***
(Standard error)	–	–	(1.5)
Percent working in the week prior to survey interview	74.0	73.5	-0.4
(Standard error)	–	–	(2.1)
Sample Sizes			
NDNH	1,638	858	2,496
Survey	1,301	638	1,939

SOURCE: Outcomes data are based on quarterly wage records obtained in a match to the National Directory of New Hires and on the PACE 18-month follow-up survey.

NOTES: Statistically significant in a one-tailed test: * at the 90% level, ** at the 95% level, *** at the 99% level. 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 sixth and seventh follow-up quarters.

Appendix F: Treatment of Outliers

The team 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 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 . (Trimming by definition creates item nonresponse since 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.” Since there is no known way to remove bias due to non-ignorable nonresponse, trimming is likely to create uncorrectable biases in estimated treatment effects.)

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 (1996) and Lumley et al. (2002) indicate that, for the sample sizes available in this evaluation, OLS (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) and credits. The evaluation team found no values that were clearly impossible and thus retained all reported values in the analysis.

Appendix G: Quarterly Impact Detail for Chapter 6 Exhibits

Exhibit G-1: Impacts on Earnings and Employment in Successive Follow-up Quarters

Outcome	Treatment Group	Control Group	Impact (Difference)	Standard Error	p-Value
Average Earnings in Follow-up Quarter:					
-2	1,663	1,628	35	53	0.257
-1	1,782	1,718	65	67	0.166
0	1,441	1,741	-300 ***	54	<.001
1	900	2,263	-1,362 ***	74	<.001
2	747	2,421	-1,674 ***	83	<.001
3	765	2,766	-2,001 ***	96	<.001
4	2,562	2,964	-402 ***	112	<.001
5	4,966	3,172	1,794 ***	140	<.001
6	5,381	3,375	2,006 ***	147	<.001
7	5,527	3,744	1,783 ***	160	<.001
8	5,601	3,898	1,703 ***	165	<.001
9	6,126	4,344	1,781 ***	176	<.001
10	6,266	4,527	1,739 ***	180	<.001
11	6,419	4,631	1,788 ***	183	<.001
Percent Employed in Follow-up Quarter:					
-2	60.3	58.5	1.8	1.5	0.115
-1	62.9	59.3	3.5 **	1.7	0.016
0	55.3	65.2	-9.8 ***	1.5	<.001
1	42.2	71.1	-28.9 ***	1.7	<.001
2	35.0	71.7	-36.7 ***	1.8	<.001
3	32.4	73.0	-40.5 ***	1.8	<.001
4	66.7	74.7	-8.0 ***	1.8	<.001
5	79.3	75.3	4.0 ***	1.6	0.007
6	81.4	77.4	4.0 ***	1.7	0.009
7	79.2	74.2	5.0 ***	1.7	0.002
8	79.1	76.2	2.9 **	1.7	0.046
9	79.5	77.5	2.0	1.7	0.115
10	80.8	78.9	1.9	1.6	0.118
11	81.3	81.8	-0.5	1.6	0.624
Sample size	1,638	858			

SOURCE: April 2018 match to wage records in the National Directory of New Hires for 1,638 treatment and 858 control group members.

NOTES: Statistically significant in a two-tailed test: * at the 10-percent level, ** at the 5-percent level, *** at the 1-percent level.

Exhibit G-2: Impacts on College Enrollment and Cumulative Full-time Equivalent Months of Enrollment in Successive Follow-up Quarters

Outcome	Treatment Group	Control Group	Impact (Difference)	Standard Error	p-Value
Percent Enrolled in College in Follow-up Quarter:					
-2	22.0	19.9	2.1	1.6	0.173
-1	21.3	18.8	2.5	1.5	0.105
0	47.1	16.2	30.9***	1.6	<.001
1	59.1	18.3	40.8***	1.7	<.001
2	55.4	25.5	29.9***	1.8	<.001
3	47.0	25.8	21.1***	1.9	<.001
4	16.4	25.8	-9.4***	1.7	<.001
5	14.4	25.5	-11.1***	1.7	<.001
6	16.9	24.7	-7.8***	1.6	<.001
7	17.2	23.7	-6.4***	1.6	<.001
8	18.7	19.6	-1.0	1.6	0.544
9	19.1	19.9	-0.8	1.6	0.621
10	17.5	19.3	-1.8	1.6	0.248
Average Cumulative Number of Full-Time Equivalent Months Enrolled in College from Month of Random through End of Follow-up Quarter:					
0	0.7	0.2	0.5***	0.0	<.001
1	2.2	0.5	1.7***	0.1	<.001
2	2.9	1.0	2.0***	0.1	<.001
3	3.5	1.4	2.1***	0.1	<.001
4	3.7	1.9	1.8***	0.1	<.001
5	4.0	2.4	1.6***	0.2	<.001
6	4.3	2.9	1.4***	0.2	<.001
7	4.6	3.3	1.3***	0.2	<.001
8	4.9	3.7	1.2***	0.2	<.001
9	5.3	4.1	1.2***	0.2	<.001
10	5.6	4.5	1.1***	0.3	<.001
Sample size	1,668	871			

SOURCE: April 2017 match to college records in the National Student Clearing House for 1,668 treatment and 871 control group members.

NOTES: Statistically significant in a two-tailed test: * at the 10-percent level, ** at the 5-percent level, *** at the 1% level.