Madison Area Technical College Patient Care Pathway Program: Implementation and Early Impact Report

Appendices

Pathways for Advancing Careers and Education

OPRE Report No. 2018-48

September 2018

PACE
Pathways for Advancing Careers and Education

Madison Area Technical College Patient Care Pathway Program: Implementation and Early Impact Report

Pathways for Advancing Careers and Education (PACE)

OPRE Report No. 2018-48

September 2018

Rachel Cook, Jill Hamadyk, Matthew Zeidenberg, Howard Rolston, and Karen Gardiner, 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.
6130 Executive Boulevard
Rockville, MD 20852

Contents

List of Exhibits

Appendix A: Patient Care Pathway Program Eligibility Criteria

To be eligible for the Patient Care Pathway program, a student had to (1) express an interest in a healthcare diploma or degree program and (2) fall within the designated Compass score range or have completed equivalent coursework at Madison College. The specific Compass score and coursework eligibility requirements that the program developed for each Patient Care Pathway academy are listed in the boxes below.

Patient Care Academy 1 (PCA1)

1. A student may qualify based on COMPASS scores alone if he or she meets one or more of the following:
   - COMPASS eligibility scores published by the college: Pre-Algebra 30-55 OR Writing 31-70 OR and Reading 61-80; OR
   - The "unpublished" eligibility scores: Pre-Algebra 27-55 OR Writing 30-70 OR and Reading 59-80; OR
   - If one or two scores fall into the published or unpublished range and the remaining score(s) are higher than the range (i.e., one or two scores prevent the student from enrolling in course he/she needs). For example, a student with the following COMPASS scores *would be* eligible: Pre-Algebra 40, Writing 40, Reading 95.
     - A student is *not* eligible if one or more scores fall below the unpublished range without any coursework (even if the other scores are within the range). For example, a student with the following COMPASS scores *would NOT* be eligible: Pre-Algebra 32, Writing 21, Reading 62 unless they had successfully completed writing coursework equivalent to the required COMPASS score range (see below).
2. If a student's COMPASS scores are below the unpublished eligibility range, they may still be eligible if they have successfully completed the following coursework:
   - Math: student has completed Math Concepts (equivalent to a Pre-Algebra 30) or Basic Algebra (equivalent to a Pre-Algebra 43)
   - Reading: student has completed Academic Reading 3&4 (equivalent to a Reading 51), Academic Reading 5&6 (equivalent to a Reading 65), College Reading Strategies (equivalent to a Reading 80), Academic Writing 3&4 (equivalent to a Reading 51), Academic Writing 5&6 (equivalent to a Reading 65), or Intro to College Writing (equivalent to a Reading 80)
   - Writing: student has completed Academic Writing 3&4 (equivalent to a Writing 31), Academic Writing 5&6 (equivalent to a Writing 46), or Intro to College Writing (equivalent to a Writing 69)
   - Note: if one of the students' scores is below the COMPASS range and they do not have coursework, they are *ineligible* even if they have coursework to make up for a low score in another subject.
3. Student may qualify if they failed or withdrew from Body Structure and Function or Medical Terminology if:
   - All COMPASS scores are within or above the range (or they have completed equivalent coursework), but student failed or withdrew from the Medical Terminology or Body Structure and Function course.

Patient Care Academy 2 (PCA2)

1. A student may qualify based on COMPASS scores alone if he or she meets one or more of the following:
   - Meets COMPASS eligibility scores published by the college: Pre-Algebra 30-Algebra 30 OR Writing 70+ OR Reading 80+; OR
   - Meets the "unpublished" eligibility scores: Pre-Algebra 27+ OR Writing 46+ with an E-Write score of 6 OR Writing 65+ with an E-Write score of 5 OR Reading 75+; OR
   - One or two scores fall into the published or unpublished range and the remaining scores are higher than the range (i.e., one or two scores prevent the student from enrolling in course he/she needs). For example, a student would be eligible if they had the following COMPASS scores: Pre-Algebra 80, Writing 81, Reading 99.
     - A student is *not* eligible if one or more scores fall below the unpublished range without any compensating e-write scores or coursework. For example, a student with the following COMPASS scores *would NOT* be eligible for PCA2 without additional coursework (see below): Pre-Algebra 80, Writing 81, Reading 60.
2. If a student's COMPASS scores are below the unpublished eligibility range, they may still be eligible if they have successfully completed the following coursework:
   - Math: student has completed Math Concepts (equivalent to a Pre-Algebra 30) or Basic Algebra (equivalent to a Pre-Algebra 43)
   - Reading: student has completed College Reading Strategies (equivalent to a Reading 80), Intro to College Writing (equivalent to a Reading 80), Academic Writing 5/6 or Written Communication (equivalent to a Reading 99+)
   - Writing: student has completed Intro to College Writing (equivalent to a Writing 69)Academic Writing 5/6 or Written Communication (equivalent to a Writing 69)
   - Note: if one of the students' scores is below the COMPASS range and they do not have coursework, they are *ineligible* even if they have coursework to make up for a low score in another subject.
3. A student may qualify if they failed or withdrew from General Chemistry if:
   - All COMPASS scores are within or above the range (or they have completed equivalent coursework), but student failed or withdrew from the General Chemistry course.

Patient Care Nursing Assistant (PCNA)

1. Students may qualify on COMPASS score alone if:
   - They have COMPASS Reading score between 61 and 80
2. If the student's COMPASS scores are below the range, they may still be eligible if they have successfully completed the following coursework
   - Reading: Academic Reading 5/6 or College Reading Strategies
3. Students may qualify if they got a D, failed or withdrew from a CNA course before
   - If the Reading COMPASS scores are in the range but the student was not successful (D, F or W) in CNA, they are eligible.

Appendix B: Baseline Characteristics and Adjustments

This appendix describes specifications for baseline covariates—including the approach to missing values in Section B.1. It then compares distributions for treatment and control group members on these measures (B.2). Finally, Section B.3 explains how the analyses control for these covariates in estimating impacts.

B.1 Details on Baseline Covariates

Exhibit B-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).

The research team imputed values for missing covariates using SUDAAN/IMPUTE, a weighted hot-deck imputation procedure (Research Triangle Institute 2012). This imputation step 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 the following: treatment-control status, site, National Student Clearinghouse (NSC)-reported enrollment status (some or none), NSC-reported credential award (some or none), and number of months of NSC-reported enrollment.[1]

Exhibit B-1. Operationalization of Baseline Measures Used as Covariates in Regression-Adjusted Impact Estimates

Variable Description	Operationalization Details	Data Source(s) Instrument & Item Number
Age	Categorical measure: Under 21 21-24 25-34 35+*	BIF: B2_dob RABIT: R_RA_Date_Assigned
Female	Binary variable 1 if female 0 if male	BIF: B7

[1] 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.

Variable Description	Operationalization Details	Data Source(s) Instrument & Item Number
Race-ethnicity	Categorical measure: Any race, Hispanic Black, non-Hispanic White, non-Hispanic* Other, non-Hispanic	BIF: B9
Family structure	Categorical measure: Spouse/partner, with children Spouse/partner, without children Single, with children* Single, without 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 own parents	Binary variable: 1 if living with own parent(s) 0 otherwise (Presence of parents of spouse not considered)	BIF: B13
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*	BIF: B23
Highest level of education completed	Categorical measure: No college* Under 1 year's college credit 1 year+ of college credit Associate's degree or above	BIF: B17
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
Academic discipline[a]	Average of 10 items (scale ranging 1-6) after reversing responses to negatively phrased items. Missing if seven or more of 10 responses blank	SAQ: S11a
Training commitment[a]	Average of 10 items (scale ranging 1-6) after reversing responses to negatively phrased items. Missing if seven or more of 10 responses blank	SAQ: S11b
Academic confidence[a]	Average of 12 items (scale ranging 1-6) after reversing responses to negatively phrased items. Missing if nine or more of 12 responses blank	SAQ: S11d
Emotional stability[a]	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

Variable Description	Operationalization Details	Data Source(s) Instrument & Item Number
Family income in past 12 months	Categorical measure: Less than $15,000 $15,000-29,999 $30,000+*	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 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
Expecting to attend school part-time if accepted	Binary variable: 1 if yes 0 if no	SAQ: S1
Frequency of situations interfering with school, work, job search, or family responsibilities	Average of six items of frequency of problems (scale ranging 1-5). Missing if four or more of six responses blank	SAQ: S15
Stress[b]	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

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.

[a] 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).

[b] Cohen et al. (1983).

B.2 Comparing Treatment and Control Groups at Baseline

Exhibit B.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 it is said that "baseline balance" was achieved. The list expands somewhat on the characteristics in Chapter 2, Exhibit 2-2.

The last column contains *p*-values for tests of hypotheses that no systematic differences exist between the treatment and control groups. On average, one would expect that out of 28 tests three will fall outside a 90-percent confidence interval due to chance. In this case, there were three statistically significant differences, which are highlighted in red. The team carefully reviewed data processing and other operations but could find no causes for these differences. It is likely that these are simply random results. Furthermore, as described in the next section, regression adjustment helps to control for any effects that chance differences might have on the impact estimates.

Exhibit B-2. Baseline Balance for Patient Care Pathways Program

Characteristic	All Participants	Treatment Group	Control Group	p-Value
Age (%)				.238
20 or under	23.4	19.9	26.9	
21 to 24	21.2	23.1	19.3	
25 to 34	30.0	32.3	27.7	
35 or older	25.4	24.7	26.1	
Female (%)	84.3	86.4	82.3	.203
Race/Ethnicity (%)				.021
Any race, Hispanic	8.8	12.4	5.3	
Black, Non-Hispanic	20.8	19.8	21.8	
White, Non-Hispanic	67.3	64.2	70.4	
Other, Non-Hispanic	6.4	7.8	4.9	
Family Structure (%)				.530
Not Living with Spouse/Partner and not Living with Children	43.5	40.7	46.5	
Not Living with Spouse/Partner but Living with Children	18.3	20.3	16.2	
Living with Spouse/Partner and not Living with Children	19.7	19.9	19.5	
Living with Spouse/Partner and Children	18.5	19.1	17.8	
Living with Parents (%)	26.5	22.8	30.3	.060
One Parent Has at Least some College (%)	51.8	50.4	53.2	.542
High School Grades (%)				.255
Mostly Got A's	6.1	6.3	5.9	
Mostly Got B's	43.1	46.7	39.5	
Mostly Got C's or Below	50.8	47.1	54.6	

Characteristic	All Participants	Treatment Group	Control Group	p-Value
Current Education (%)				.211
Less Than a High School Degree	3.0	4.8	1.2	
High School or Equivalent	44.4	44.8	43.9	
Less Than 1 Year of College	24.8	24.4	25.2	
1 or More Years of College	21.6	20.4	22.8	
Associate's Degree or Higher	6.3	5.6	6.9	
Received Vocational or Technical Certificate or Diploma (%)	39.5	41.1	38.0	.481
Career Knowledge Index (mean)	0.41	0.43	0.39	.196
Psycho-Social Indices (means)				
Academic Discipline Index	5.06	5.05	5.07	.786
Training Commitment Index	5.59	5.61	5.58	.628
Academic Self-Confidence Index	4.43	4.46	4.39	.361
Emotional Stability Index	4.99	5.04	4.95	.203
Social Support Index	3.31	3.30	3.32	.493
Stress Index	2.22	2.21	2.24	.748
Depression Index	1.55	1.52	1.57	.310
Family Income (%)				.038
Less than $15,000	25.6	27.6	23.5	
$15,000-$29,999	29.9	33.5	26.1	
$30,000 or More	44.6	38.9	50.4	
Family Income (mean)	$33,165	$31,694	$34,694	.280
Public Assistance/Hardship Past 12 Months				
Received WIC or SNAP (%)	35.6	32.9	38.2	.193
Received Public Assistance or Welfare (%)	4.4	5.1	3.7	.600
Reported Financial Hardship (%)	34.3	35.5	33.2	.528
Current Work Hours (%)				.949
0	27.9	27.4	28.5	
1 to 19	11.5	11.7	11.4	
20 to 34	32.6	31.9	33.3	
35 or more	27.9	29.0	26.8	
Expected Work Hours in Next Few Months (%)				.227
0	18.3	16.4	20.3	
1 to 19	15.1	18.1	12.1	
20 to 34	47.4	47.8	47.0	
35 or more	19.2	17.7	20.7	
Life Challenges Index (mean)	1.44	1.45	1.43	.710
Owns a Car (%)	84.9	86.7	83.1	.207
Has both Computer and Internet at Home (%)	84.1	84.3	83.9	.900
Ever arrested (%)	20.7	23.6	17.8	.139
Sample Sizes	499	250	249	

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

NOTES: Tests for statistically significant imbalance were based on SAS/TTEST procedure.

B.3 Regression Adjustment

In this appendix, the team describes the regression adjustment approach used to improve precision and minimize effects of sampling error on impact point estimates.

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

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

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, $\text{var}\left(\hat{\delta}\right) \approx \left(1-R^2\right)\text{var}(\bar{y}_T - \bar{y}_C)$, where R^2 is proportion of the variance in Y_i that can be explained by X_i, in Equation B.2 below.

The team's analyses showed that the method can perform poorly when the number of baseline covariates is relatively large compared with the number of observations. Specifically, when the ratio *n/p* is not very large, it can happen that $\text{var}\left(\hat{\delta}\right) > \text{var}(\bar{y}_T - \bar{y}_C)$, meaning that the variance on the estimated treatment effect using the regression adjustment in Equation B.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.[2] There is a lack of good research on how large the ratio of cases to variables needs to be in order to guarantee that $\text{var}\left(\hat{\delta}\right) < \text{var}(\bar{y}_T - \bar{y}_C)$, 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 2016).

[2] For example, with a sample size of 1,000, 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).

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 B.2 below). The next step was to calculate how different each program and control group member's outcome was from what would have been expected under control conditions, as in Equation B.3. These differences between actual and predicted outcomes are called "residuals." The team then calculated the difference between the average residual in the program group and the average residual in the control group, as in Equation B.4.

Equation B.5 gives the formula used to estimate standard errors on these impact estimates.

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

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

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

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

For survey-based outcomes subject to nonresponse, the team used a weighted version of this estimator, as in Equation B.6.

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

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

This method is similar to the method developed by Koch and coauthors (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 coauthors fit Equation B.2 on the entire sample rather than just on the control sample. The main advantage of fitting Equation B.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 Equation B.2 on just the control sample will increase the estimated standard errors obtained in Equation B.5 compared with 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 improved precision relative to both pure randomization and OLS (see Equation B.1). The variance on the estimate of the impact of the program on the primary outcome (credits earned through month 18) was nine percent smaller with the modified Koch's estimator than it would have been with the OLS approach. Across a collection of primary and secondary outcomes, the average variance reduction due to using the modified Koch's estimator instead of the OLS estimator was 11 percent.

Exhibit B-3 shows the regression coefficients from Equation B.2 for the confirmatory outcome, credits earned through 18 months. 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. Note that of the 34 baseline covariates allowed into the model, only three of these are predictive of future credits for the control group sample. Specifically, low emotional stability, plans to work 40 or more hours per week in addition to going to school if admitted into the program, and plans to only attend school part-time are negatively associated with future credits.

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 study.

Exhibit B-4 shows impacts on selected confirmatory and secondary outcomes before and after regression adjustment. The similarity in estimates further attests to the high level of balance achieved through random assignment. There are almost no changes in the characterization of the strength of evidence against the null hypothesis (the "star pattern") and the only estimates that appear to change very much were subject to high standard errors.

Exhibit B-3. Coefficients for Baseline Characteristics as Predictors of College Credits Earned through 18 Months: Patient Care Pathways Program Control Group Members

Baseline Covariate	Estimate	Standard Error	*p*-Value
Intercept	18.15	15.59	.245
Age			
20 or under	–2.54	3.41	.457
21 to 24	–0.48	2.83	.865
25 to 34	–2.19	1.90	.251
35 or older	0	na	na
Sex			
Female	0.50	2.15	.817
Male	0	na	na
Race/Ethnicity			
Any race, Hispanic	–2.31	2.40	.338
Black, Non-Hispanic	–1.40	2.08	.501
White, Non-Hispanic	0	na	na
Other, Non-Hispanic	1.76	3.19	.582
Family Structure			
Not Living with Spouse/Partner and Not Living with Children	–1.56	2.52	.536
Not Living with Spouse/Partner but Living with Children	2.29	2.56	.372
Living with Spouse/Partner and Not Living with Children	0	na	na
Living with Spouse/Partner and Children	1.60	2.19	.465
Living with Parents	–0.15	2.97	.959
One Parent Has at Least Some College	2.12	1.54	.171
High School Grades			
Mostly Got A's	4.24	3.89	.277
Mostly Got B's	2.45	1.59	.126
Mostly got C's or Below	0	na	na
Current Education			
High School Degree or Less	0	na	na
Less Than 1 Year of College	–0.25	2.13	.907
1 or More Years of College	3.23	2.20	.143
Associate's Degree or Higher	–3.50	3.22	.277
Career Knowledge Index	1.14	2.12	.590
Family Income			
Less than $15,000	–1.06	2.08	.612
$15,000-$29,999	–0.94	2.10	.657
$30,000 or More	0	na	na
Psycho-Social Indices			
Academic Discipline Index	–2.07	1.94	.286
Training Commitment Index	1.95	2.45	.427
Academic Self-Confidence Index	2.27	1.43	.114
Emotional Stability Index	–2.96	1.27	.020
Stress Index	0.97	1.39	.484

Baseline Covariate	Estimate	Standard Error	*p*-Value
Depression Index			
Life Challenges Index	–1.93	2.32	.406
Public Assistance/Hardship Past 12 Months			
Received WIC or SNAP	–0.98	1.91	.610
Received Public Assistance or Welfare	2.52	3.34	.452
Reported Financial Hardship	–0.77	1.74	.658
Current Work Hours			
0 to 19	0	na	na
20 to 34	–0.76	1.87	.684
35 or more	0.81	2.50	.746
Expected Work Hours in Next Few Months			
0 to 19	0	na	na
20 to 34	–3.29	2.24	.143
35 or more	–5.15	2.82	.069
Plan to attend school only part-time if admitted to the Patient Care Pathways program	–2.69	1.46	.068

SOURCE: Abt Associates calculations based on data from Madison College records, the PACE Basic Information Form (BIF), and the PACE Self-Administered Questionnaire (SAQ).
NOTES: Model estimated with SAS/SURVEYREG procedure. Sample size=249.

Exhibit B-4. Comparison of Selected Impact Estimates with and without Adjustment for Baseline Imbalances

Outcome	Unadjusted Est (StdErr)	Adjusted Est (StdErr)
Regular college credits earned during months:		
1-6	0.377 (0.335)	0.315 (0.325)
7-12	0.544 (0.442)	0.382 (0.427)
13-18	0.118 (0.421)	0.148 (0.415)
1-18	1.039 (0.952)	0.844 (0.908)
Proportion enrolled in college occupational training during months:		
1-6	0.0849*** (0.0350)	0.0769** (0.0345)
7-12	0.0815** (0.0419)	0.0716** (0.0414)
13-18	0.0898** (0.0437)	0.0951** (0.0423)
1-18	0.0647** (0.0321)	0.0653** (0.0317)
Total hours of occupational training at:		
A college	16.3 (16.7)	9.8 (15.8)
Proportion who received a credential from:		
A college	0.0314 (0.0338)	0.0076 (0.0333)

SOURCE: Abt Associates calculations based on data from the administrative records and response status to the 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 in Equations (B.4) and (B.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 10-percent level.

Appendix C: College Records Data

The research team used records from Madison College to measure college outcomes in the study. Such administrative data offer strong advantages over survey data—notably in avoiding loss of sample to nonresponse and any effects on data quality from survey response errors.

The main question was whether Madison College data could support sufficiently broad measures of college enrollment given that some sample members also attended other colleges. To assess the extent of enrollment at other colleges, the team matched the sample to college records maintained in the National Student Clearinghouse (NSC). Covering 96 percent of college enrollments nationwide, the NSC data provide an excellent frame for these purposes.[3] As shown in Exhibit C-1, the vast majority of sample members enrolling in college (over the period from randomization to November 2015) enrolled only at Madison College and relatively few enrolled only at other colleges. Rates of non-host college enrollment are slightly higher in the control group.

Exhibit C-1. National Student Clearinghouse–Reported Enrollment at Madison College and Other Colleges for the Pathways to Health Care Evaluation Sample, by Study Group

	Enrollment Documented in National Student Clearinghouse				
	Enrolled at Madison College			Enrolled Only at Colleges Other Than Madison College (%)	Total Ever Enrolled (%)
Group	At Host College (%)	Only at Madison College (%)	At Madison and Other College (%)		
Treatment	99.2	85.6	13.6	0.8	100.0
Control	97.2	80.5	16.7	2.8	100.0

SOURCE: Abt Associates calculations based on data from the National Student Clearinghouse (NSC).
NOTE: NSC data cover the period from randomization (for the person) to November 2015, a period that varies in length from 23 months to 48 months, depending on how early the student was randomized.

Based on these results, the team devised an imputation approach that used NSC data to adjust Madison College data for enrollment spells[4] at non-host colleges.[5] The goal of imputation was to fill in the missing instructional hours, earned credits, and earned credentials for every spell

[3] NSC's coverage evaluation shows that coverage of for-profit colleges can be relatively low, but the for-profit sector is very small in Wisconsin and PACE survey data show negligible numbers went to for-profit schools in either arm.

[4] A spell was defined as a period of enrollment with no gaps longer than three months unless the gap included one of the summer months (June, July, and August), in which case a gap of seven months was required to initiate a new spell.

[5] Though useful for this purpose, the NSC data were not in themselves sufficient because they exclude some key study outcomes (e.g., instructional hours and credits) and do not completely cover others (e.g., certificates and degrees).

that the NSC showed the student spent at another college by pairing every other-college spell with a similar Madison College spell, and then copying the information over.

The imputation strategy involved several steps:

1. Find a Madison College record for as many NSC-reported Madison College spells as possible. This step filled in instructional hours, earned credits, and earned credentials for most NSC-reported Madison College spells. The team referred to this step as the *exact matching* process because there is a single correct match in the Madison College system for almost all the NSC-reported Madison College spells.
2. Resolve NSC-reported spells that did not match to a Madison College record. For such NSC-reported spells, the team assumed that no instructional hours were received, no credits were earned, and no credentials were earned.
3. Summarize the available data for each NSC-reported spell and the student to whom the spell belonged. The team summarized these data by developing statistical models that predict four critical Madison College–reported outcomes for each spell.[6]
4. Match each NSC-reported other-college spell with a "similar" NSC-reported Madison College spell in terms of the predicted four critical outcomes. The team referred to this step as *statistical matching* because there are many possible NSC-reported Madison College spells that could be matched to every NSC-reported other-college spell. The team only matched other-college spells of students in the treatment group to Madison College spells or other students in the treatment group. A parallel restriction was placed on the matching of other-college spells of students in the control group. The team imposed these restrictions to avoid "washing out" any effects by making control experiences artificially more similar to treatment experiences.
5. Lastly, copy the information from steps 1 and 2 that were associated with every NSC-reported Madison College spell over to the statistically matched NSC-reported other-college spell.

The following sections give more information for each step.

C.1 Details on Step 1 (Exact Matching)

The team conducted the exact matching of each NSC-reported spell at Madison College with a Madison College–reported spell by determining the amount of overlap between the spells, based on the start and end dates of each spell. If only one Madison College–reported spell overlapped with an NSC-reported spell at Madison College, then the team considered the two

[6] Each of the four predicted outcomes is a linear function of the larger collection of available covariates, and therefore "summarizes" the larger collection.

spells to be matched without regard to how well start and end dates aligned between the two systems. If multiple Madison College–reported spells overlapped with one NSC-reported spell at Madison College, then the team considered the Madison College–reported spell with the most months of overlap to be matched to the NSC-reported spell. If one Madison College–reported spell overlapped with multiple NSC-reported spells at Madison College, then the Madison College–reported spell was broken into pieces that better matched the NSC-reported spells. The team then transcribed the training hours, credits, and credentials associated with the Madison College–reported spell in the Madison College record system over to the NSC-reported spell.

C.2 Details on Step 2 (Unmatched NSC-Reported Madison College Records)

In a small number of cases, NSC-reported spells at Madison College did not overlap with any Madison College–reported spells. Based on data investigation with Madison College staff, it was determined that many of these cases are due to early course drops or withdrawals. Because the courses were dropped early in the term, they were not included in the file Madison College provided to the research team. However, the college does appear to include these records in extracts it sends to the NSC. For these spells, the team assumed that the student had experienced zero hours of instructional credits and earned no credits or credentials.

C.3 Details on Step 3 (Data Summarization)

The available data about each spell that could be used included NSC-reported spell duration and timing, NSC-reported credentials awarded in connection with the spell, and self-reported baseline variables. Follow-up survey data could not be used because the team wanted to do the matching both on survey respondents and on nonrespondents. To facilitate matching, the team developed statistical models for four Madison College–reported outcomes on the set of exactly matched records in terms of these variables. The Madison College–reported outcomes were:

- Credits earned within the intersection of the spell (as reported by Madison College) with the 12-month window after randomization
- Credits earned within the intersection of the spell (as reported by Madison College) with the 18-month window after randomization
- Hours of non-credit occupational training within the intersection of the spell (as reported by Madison College) with the 12-month window after randomization
- Hours of non-credit occupational training within the intersection of the spell (as reported by Madison College) with the 18-month window after randomization

These Madison College–reported outcomes jointly define the primary outcome and some secondary outcomes.

The procedure involved first fitting models for NSC-reported Madison College spells and then using estimated coefficients to predict values for both Madison College spells and spells at other colleges. The models (not shown) involved 12 to 18 characteristics from NSC and from the baseline.

C.4 Details on Step 4 (Statistical Matching)

For each spell at a college other than Madison College, the team calculated the weighted Euclidean distance from that spell to every spell at Madison College as:

$$D_{ij} = \sqrt{\sum_{\ell=1}^{4} c_{\ell} \left(\hat{z}_{\ell i} - \hat{z}_{\ell j} \right)^2} \qquad \text{(C.1)}$$

where $\hat{z}_{1i}, \ldots \hat{z}_{4i}$ are the predicted outcomes for the spell at a college other than Madison College, $\hat{z}_{1j}, \ldots \hat{z}_{4j}$ are the predicted outcomes for a spell at Madison College, and $c_1, \ldots, c_4$ are emphasis weights.[7] The team selected the Madison College spell *j* within the same study group (treatment/control) that minimized D_{ij} as the matched Madison College spell for the *i*-th other-college spell. The same Madison College spell could be matched to more than one other-college spell. The procedure is known in the literature as "predictive mean matching" (van Buren 2012).

C.5 Details on Step 5 (Propagating Madison College Values)

The final step entailed copying matched data on instructional hours, earned credits, and earned credentials from Madison College records to serve as values for a spell at the other college. The procedure involved both the exact matching and the statistical matching. The outcomes of an NSC-reported spell at another college were copied over from Madison College–reported outcomes of the Madison College–reported spell that had been exactly matched to the NSC-reported Madison College spell that had been statistically matched to the other-college spell. The team did this separately for every outcome based on Madison College records, including hours, credits, and credentials within six-month anniversaries of the randomization date for the person.

Exhibit C-2 provides summary statistics on selected outcomes reflecting the quality of resulting imputations. Imputation involved a total of 50 NSC-reported spells at colleges other than Madison College. The team found Madison College matches for each of the 50 spells from the 441 NSC-reported Madison College spells. The matches were generally of high quality. For

[7] The team gave slightly larger weights to the 18-month predictions.

example, in the control group, the correlation in predicted earned credits through 18 months across matched pairs was 0.99. In the treatment group, it was only slightly lower at 0.98.

Exhibit C-2. Descriptive Statistics Related to Imputation of Credits and Hours for Non-Host College Spells

Statistic	Control	Treatment
Number of spells at colleges other than Madison College	31	19
Number of spells at Madison College	203	238
Correlation across matched pairs (other-college spell and Madison College spell) in predicted:		
Hours of non-credit occupational training hours through 12 months	0.74	0.92
Earned credits through 12 months	0.98	0.99
Hours of non-credit occupational training hours through 18 months	0.97	0.83
Earned credits through 18 months between spell at Madison College and matched spell at some other college	0.99	0.98

SOURCE: Abt Associates calculations based on data from the National Student Clearinghouse, Madison College records, the PACE Basic Information Form (BIF), the PACE Self-Administered Questionnaire (SAQ), and the PACE short-term follow-up survey.

Appendix D: Survey Data Recoding and Adjustments

This appendix documents key technical detail for impact estimates for outcomes based on 18-month follow-up survey data. Section D.1 documents coding for scales based on follow-up survey data. Section D.2 describes the imputation process for some missing survey data elements. Section D.3 analyzes survey nonresponse and documents the decision not to apply nonresponse weights in the impact analysis.

D.1 Measures Based on Follow-up Survey Data

Exhibit D-1 provides details on specifications for the process outcomes analyzed in the Implementation Analysis of Chapter 4.

Chapter 5, Exhibit 5-1 provided descriptions of outcomes in the impact analysis of the Patient Care Pathways Program. Exhibit D-2 provides details on the operationalization of each measure and the underlying survey questions.[8]

Exhibit D-1. Details on Specifications for Survey-Based Outcomes in Chapter 4

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Receipt of Education or Training		
Entire Study Sample		
Received education or training since random assignment		
In any subject/field	Two-question format with slightly different wordings to try to get all training spells reported	A1, A1a
In a healthcare occupation	Open-ended responses about name of target occupation and understanding of future duties were coded by staff from the U.S. Census Bureau into the U.S. Department of Labor's Standard Occupational Classification (SOC) codes. Those in programs designed to train them for jobs as Healthcare Practitioners/Technicians (SOC 29-xxxx) or Healthcare Support Workers (SOC 31-xxxx) were counted for this outcome. This does not include office workers in the healthcare industry or personal care aides in nursing homes	A19a, A20, A21, A27a, A27c, A27d
Since random assignment, ever attended	The team looked up place names reported in question A4 in the Integrated Postsecondary Education Data System and used the IPEDS classification to edit self-reports in question A5. Private for-profit colleges were not counted as proprietary schools. Only places not classified as degree-granting in IPEDS and that are privately run for profit were classified as proprietary schools	A4, A5

[8] This table applies to all nine PACE sites. For the evaluation of the Patient Care Pathways program, the team substituted parallel outcomes based on administrative data wherever possible. For some variables such as receipt of a credential from any source, that meant combining administrative and survey data. For these hybrid recodes, the sample size is just the survey respondent sample size.

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Two-year college	Community or technical college (two-year college)	
Four-year college	Four-year college/university	
Proprietary school	Private school/company that provides training	
Adult high school/education	Adult education / adult high school / community school / night school	
Community/non-profit organization		
Other	State unemployment/employment office, One-Stop career center, your place of employment, or somewhere else	
Of Those Who Attended Any Education or Training		
Time spent at school and work at first place attended	Question was asked about each place attended since randomization, but only information on first place was analyzed. Enrollment dates were used to determine first place attended since randomization	A7
Full-time school and full-time work		
Full-time school with no or part-time work		
Part-time school and full-time work		
Part-time school with no or part-time work		
Views of classes at first place attended	Questions about career relevance and learning methods were asked only about first place attended. This was done to reduce respondent burden. First place was chosen rather than last place because PACE programs put particular emphasis on innovative teaching methods for basic education classes, which would typically be the first classes taken	
Strongly agrees relevant to life/career[a]	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
Used active learning methods most/all of the time[b]	Responses to six-item battery were reverse scaled (1=none of the time, 4=all the time) and then averaged. Anyone with an average of 2.5 or larger was counted	A47a-A47f
Perceived strong emphasis on community	People who responded "a great deal" were counted	A37
Basic Skills Instruction and Tests		
Received basic skills instruction since random assignment		
Academic skills		A10b
English as a Second Language		A10a
Took college placement exam		
English		A57
Math		A58

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Passed college placement exam		
English		A57a
Math		A58a
Life Skills Instruction		
Received life skills instruction since random assignment		A10e
Receipt of Various Supports		
Received assistance from any organization since random assignment (%)	This was asked of everyone, even those with no training since randomization	A62
Career counseling		
Help arranging supports for school/work/family		
Job search or placement		
Cited financial support as challenge in enrollment or persistence[b]	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
Received supports at first place of instruction attended (%)	Question was asked about first and second places attended since randomization, but only information on first place was analyzed. Enrollment dates were used to determine first place attended since randomization	
Career counseling		A36d
Ever		
Three or more times		
Academic advising		A36a
Ever		
Three or more times		
Financial advising		A36b
Ever		
Three or more times		
Tutoring		A36d
Ever		
Three or more times		
Help arranging supports for school or work		A36f
Ever		
Three or more times		
Job search/placement assistance		A36e
Ever		
Three or more times		

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Received financial assistance at first place of instruction (%)[a]	Question was asked about each place attended since randomization, but only information on first place was analyzed. Enrollment dates were used to determine first place attended since randomization	
Grants/scholarship	A Pell grant or other government grant or scholarship—not counting loans you have to pay back, Must indicate in question A31 that funds were used for tuition, other school-related expense, or living expenses	A30g, A31
Loan	Loans in your own name or loans in your parents' names. Must indicate in question A31 that funds were used for tuition, other school-related expense, or living expenses	A30e, A30f
Offered opportunities for related work experience as part of training at first place of instruction (%)	Question was asked about each place attended since randomization, but only information on first place was analyzed. Enrollment dates were used to determine first place attended since randomization	
Clinical internship		A38b
Visits to local employer		A38c
Work-study job		A38a
Apprenticeship		A38e
Any related work experience (including other)		A38f

Exhibit D-2. Details on Specifications for Survey-Based Outcomes in Chapter 5

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Confirmatory and Secondary		
Education		
Hours of occupational training at colleges[a]	1) Students receiving non-credit occupational training were asked for duration of training (e.g., weeks) and intensity (e.g., hours per week). These were multiplied together to obtain hours of occupational training 2) If students reported earning college credits at colleges, the research team translated credits for hours using a rule of 15 hours of training time per credit. (Typical three-credit college courses at most U.S. colleges and universities meet three hours per week for 15 weeks, so each credit represents 15 hours of class time) 3) If students reported receiving both non-credit and credit training at a college, the team summed the hours from both	A24, A28, A29
Hours of occupational training at places other than colleges	Same as at colleges	A24, A28, A29
Hours of occupational training at any place	Sum of prior two outcomes	A24, A28, A29

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Confirmatory and Secondary		
Credential receipt from colleges[a]	The survey had separate questions about credentials awarded for regular credit-bearing courses and for non-credit occupational courses. It 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, Adult Basic Education, or life-skills courses	A22, A23, A27e, A27f
Credential receipt from another type of education-training institution	Same as at colleges	A22, A23, A27e, A27f
Credential receipt from a licensing/certification body	The survey asked about the highest level of occupational 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 randomization or later, then the person was coded as having earned such a credential	A56
Received a credential from any source[b]	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
Career Progress		
Employment and earning $12 or per hour	Analyzed response to survey question for control group. Selected 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 programs such as the Patient Care Pathway program	E2
Employment in job requiring 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 U.S. Department of Labor's Standard Occupational Classification (SOC) codes. The team then looked up the Job Zone[9] for each SOC code in the Bureau of Labor Standards O*NET system.[10] 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 medium preparation—seemed a reasonable goal for graduates of the program. 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

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

10 https://www.onetonline.org/ [last accessed September 12, 2016]

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Confirmatory and Secondary		
Working in a healthcare occupation	Three open-ended questions about the kind of work, usual activities, and job title were coded into one of the SOC codes. 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. Note, being employed in a healthcare occupation is usually associated with employment in the healthcare industry, but this is not always true. The survey did not ask about industry of employer	E3, E4, E5
Perceived career progress	This was a new scale created for PACE. 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 an individual's sense of progress in a career pathways program as described in Fein (2012)	C5, C6
Confidence in career knowledge	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). Items d-f were adapted from the Career Exploration Survey. Two items (c and g) were new and written specifically for the PACE BIF. Response categories range from 1=strongly disagree to 4=strongly agree	C3
Access to career supports	This was a new scale created for PACE. It is a six-item scale, counting number of types of career-supportive relationships in workforce and education settings; response categories range from 1= no to 2=yes. 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 Sorenson 2010)	C2
Exploratory		
Psycho-Social Skills		
Grit	Existing scale from Duckworth et al. (2007). The eight-item scale captures persistence and determination; response categories range from 1=strongly disagree to 4=strongly agree	B3
Academic self-confidence	Existing scale from Le et al. (2005). This scale was used for a second time in the follow-up survey. It was used initially in the BIF. The 12-item scale includes response categories that range from 1=strongly disagree to 6=strongly agree	B4
Core self-evaluation	Existing scale from Judge (2009). The 12-item scale response categories range from 1=strongly disagree to 4=strongly agree	B6
Social belonging in school	Shorter version of an existing scale by Walton and Cohen (2007, 2011). The five-item scale captured sense of belonging; response categories range from 1=strongly disagree to 4=strongly agree	B7
Life Stressors		
Financial hardship	This was a new scale created for PACE. This scale was used for a second time in the follow-up survey. It was used initially in the BIF. The two-item scale asked about financial hardship, reported as either an inability to pay rent/mortgage or not enough money to make ends meet. Response categories were 0=no or 1=yes	D1, D2

Outcome	Details on Derivation of Outcome	Follow-Up Survey Question(s)
Confirmatory and Secondary		
Life challenges	This was a new scale created for PACE. It was adapted from a longer instrument by Kessler et al. (1998). This scale was used for a second time in the follow-up survey. It was used initially in the BIF. The seven-item scale captured life challenges that interfered with school, work, or family responsibilities. The response categories range from 1=never to 5=very often	D3
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 the BIF. The four-item scale captured perceived stress. The response categories range from 1=never to 4=very often	D4

[a] Not used in the evaluation of the Patient Care Pathways program. Administrative data used instead.

[b] Hybrid of survey and administrative data used in the evaluation of the Patient Care Pathways program.

D.2 Imputation of Some Item Nonresponse in the Follow-up Survey

This section documents the research 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 did not answer subsequent questions on the nature of these experiences. 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 the National Student Clearinghouse confirmed education and training receipt and suggested misunderstanding survey questions as a likely source of the missing data. The discrepancy affected fewer than 10 percent of respondents and occurred at similar rates for treatment (nine percent) and control (seven percent) group members. Specifically, the missing data involved responses to a filter question (A10) ascertaining participation in each of a series of types of education and training activities (ESL, Adult Basic Education, courses for college credit, non-credit occupational training, life-skills classes).

To adjust for these missing data, the team imputed new responses for A10 using a nearest neighbor hot deck procedure (Andridge & Little, 2010).[11] 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

[11] 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.

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 Madison College, as well two other PACE sites (San Diego Workforce Partnership and Pima Community College) to boost power.[12] The final imputation model used 12 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 those respondents 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. For training hours, however, the team imputed responses for each type of course at each school the respondent attended. This imputation allowed the team to sum training hours across schools and types of courses 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 B.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 Survey Nonresponse Analysis

The 18-month follow-up survey obtained very similar response rates in the treatment (72 percent) and control (69 percent) groups. Such similarity suggests that any differences in baseline characteristics and outcomes for respondents and the full sample will tend to be small and similar in size for the two groups. This section assesses the implications of nonresponse for the study's impact findings.

Exhibit D-3 compares distributions on baseline characteristics for all sample members and survey respondents. Nonresponse increased the number of significant imbalances across the two arms (using a threshold of .10 for statistical significance) from three to seven statistically significant differences.

[12] Data collection was completed first at these three sites, and so they were processed together.

Exhibit D-4 compares regression-adjusted impacts on college outcomes from administrative records for the full and respondent samples. Point estimates and standard errors for impacts are generally similar, although point estimates for the survey sample tend to be somewhat smaller than those for the full sample. Such differences are not surprising given the expected sensitivity in estimates for this site's small samples.

As a further test, we 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 follow:

1. Winnow the list of potential covariates that are statistically significant in a logistic regression model for response status.[13] Do this separately for treatment and control cases. This approach identified age, financial hardship at baseline, and college persistence (per NSC records) as significant predictors of response status in the treatment sample. The team did not find any significant predictor of response status in the control arm.
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 nonrespondents.
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 did not eliminate any of the seven statistically significant baseline imbalances.[14] Furthermore, the last column in Exhibit D-4 shows that the use of weights shifts most point estimates of program impacts further from those for the full sample. For example, using the full sample, the estimated impact of the Patient Care Pathways program on credits earned through 18 months is 0.84 credits. When only the survey respondents are used, the estimated impact shrinks to 0.60 credits. If weights are used, the

---

[13] The team used the stepwise search option in SAS/LOGISTIC for this purpose with a *p*-value to enter the model of .20 and a *p*-value to stay in the model of .10.

[14] Not shown in this table, the adjustment was effective in making the weighted treatment respondent sample resemble the full treatment sample more closely. However, given that the paramount focus of this study is on treatment/control differences, the team did not believe that this improvement should be an important consideration in whether to use nonresponse adjustment weights.

estimated impact flips to a negative impact of -0.11 credits. Of the 10 administratively based outcomes analyzed in the table, weighting only improved one of the point estimates of impacts (proportion enrolled in occupational training during months 13-18 after randomization). However, even for this one, the strength of evidence against the null hypothesis was better without the weighting (two stars) than with the weighting (one star).

This relatively strong sensitivity to weighting most likely results from overfitting in nonresponse models. The likelihood of overfitting increases when the number of potential covariates is relatively large compared with treatment and control group sample sizes in this site.[15] In the team's assessment, the difficulties in estimating nonresponse models for this small sample appear to introduce more error than they remove. A close comparison of the standard errors also indicates that weighting increases the standard errors more often than it decreases them. Given these counter-indications, the team decided not to adjust for nonresponse in analyses of survey-based outcomes in this report.

[15] The team considered forming the nonresponse adjustment models on a pooled dataset involving multiple sites to overcome the sample size limitation, but was concerned that the determinates of nonresponse would vary across the sites and that the pooling would be seen as arbitrary.

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

Madison College Baseline Characteristics									
	All Participants			Survey Respondents, Unweighted			Survey Respondents, Weighted		
	Treatment	Control	*p*-Value	Treatment	Control	*p*-Value	Treatment	Control	*p*-Value
Age			0.238			0.059			.082
20 or under	19.9%	26.9%		16.6%	28.3%		18.6%	28.3%	
21 to 24	23.1%	19.3%		21.0%	18.5%		23.4%	18.5%	
25 to 34	32.3%	27.7%		32.6%	25.4%		33.5%	25.4%	
35 or older	24.7%	26.1%		29.8%	27.8%		24.5%	27.8%	
Gender			0.203			0.442			.462
Female	86.4%	82.3%		85.6%	82.6%		85.5%	82.6%	
Male	13.6%	17.7%		14.4%	17.4%		14.5%	17.4%	
Race/Ethnicity			0.021			0.078			.058
Hispanic	12.4%	5.3%		11.9%	5.3%		12.2%	5.3%	
Black non-Hispanic	19.8%	21.8%		18.1%	21.9%		19.3%	21.9%	
White non-Hispanic	64.2%	70.4%		64.4%	71.6%		62.8%	71.6%	
Other non-Hispanic	7.8%	4.9%		7.3%	4.1%		7.8%	4.1%	
Family Structure			0.530			0.547			.804
Not Living with Spouse/Partner and Not Living with Children	40.7%	46.5%		37.4%	43.7%		39.7%	43.7%	
Not Living with Spouse/Partner but Living with Children	20.3%	16.2%		20.7%	19.2%		20.8%	19.2%	
Living with Spouse/Partner and Not Living with Children	19.9%	19.5%		20.7%	21.0%		20.1%	21.0%	
Living with Spouse/Partner and Children	19.1%	17.8%		21.2%	16.2%		19.4%	16.2%	
Percent Living with Parents	22.8%	30.3%	0.060	20.1%	31.1%	0.019	21.1%	31.1%	.037
One Parent Has at Least Some College	50.4%	53.2%	0.542	48.3%	52.1%	0.483	49.3%	52.1%	.610
High School Grades			0.255			0.249			.234
Mostly Got A's	6.3%	5.9%		6.3%	6.7%		6.2%	6.7%	
Mostly Got B's	46.7%	39.5%		48.3%	39.4%		48.6%	39.4%	
Mostly got C's or Below	47.1%	54.6%		45.4%	53.9%		45.2%	53.9%	
Current Education			0.211			0.269			.293
Less Than a High School Degree	4.8%	1.2%		5.6%	1.2%		5.2%	1.2%	
High School or Equivalent	44.8%	43.9%		41.7%	42.9%		43.2%	42.9%	
Less Than 1 Year of College	24.4%	25.2%		23.3%	23.5%		23.6%	23.5%	
1 or More Years of College	20.4%	22.8%		22.8%	24.7%		22.0%	24.7%	
Associate's Degree or Higher	5.6%	6.9%		6.7%	7.7%		6.1%	7.7%	
Received Vocational or Technical Certificate or Diploma	41.1%	38.0%	0.481	46.1%	40.4%	0.281	44.0%	40.4%	.495
Career Knowledge Index (average of items)	0.43	0.39	0.196	0.44	0.38	0.132	0.43	0.38	.219

Madison College Baseline Characteristics									
	All Participants			Survey Respondents, Unweighted			Survey Respondents, Weighted		
	Treatment	Control	*p*-Value	Treatment	Control	*p*-Value	Treatment	Control	*p*-Value
Income			0.038			0.082			.064
Less than $15,000	27.6%	23.5%		25.0%	22.5%		27.8%	22.5%	
$15,000-$29,999	33.5%	26.1%		35.5%	26.3%		33.9%	26.3%	
$30,000 or More	38.9%	50.4%		39.5%	51.3%		38.3%	51.3%	
Mean	31,694	34,694	0.280	31,729	34,157	0.452	30,620	34,157	.268
Psycho-Social Indices									
Academic Discipline Index	5.05	5.07	0.786	5.09	5.06	0.616	5.07	5.06	.838
Training Commitment Index	5.61	5.58	0.628	5.62	5.56	0.215	5.62	5.56	.196
Academic Self-Confidence Index	4.46	4.39	0.361	4.52	4.38	0.133	4.49	4.38	.226
Emotional Stability Index	5.04	4.95	0.203	5.08	4.86	0.009	5.06	4.86	.018
Social Support Index	3.30	3.32	0.493	3.32	3.30	0.633	3.32	3.30	.770
Stress Index	2.21	2.24	0.748	2.15	2.25	0.201	2.17	2.25	.300
Depression Index	1.52	1.57	0.310	1.51	1.55	0.449	1.53	1.55	.625
Life Challenges Index (averages in original units 1-5)	1.45	1.43	0.710	1.42	1.45	0.544	1.44	1.45	.836
Public Assistance/Hardship Past 12 Months									
Received WIC or SNAP	32.9%	38.2%	0.193	33.5%	38.3%	0.357	35.2%	38.3%	.553
Received Public Assistance or Welfare	5.1%	3.7%	0.600	3.6%	4.8%	0.585	3.7%	4.8%	.651
Reported Financial Hardship	35.5%	33.2%	0.528	33.0%	34.9%	0.704	36.4%	34.9%	.776
Current Work Hours			0.949			0.843			.876
0	27.4%	28.5%		26.7%	28.1%		26.2%	28.1%	
1 to 19	11.7%	11.4%		12.2%	14.0%		12.4%	14.0%	
20 to 34	31.9%	33.3%		29.4%	30.4%		30.3%	30.4%	
35 or more	29.0%	26.8%		31.7%	27.5%		31.2%	27.5%	
Expected Work Hours in Next Few Months			0.227			0.233			.247
0	16.4%	20.3%		17.4%	21.7%		17.1%	21.7%	
1 to 19	18.1%	12.1%		21.0%	12.7%		20.9%	12.7%	
20 to 34	47.8%	47.0%		43.7%	45.2%		43.6%	45.2%	
35 or more	17.7%	20.7%		18.0%	20.4%		18.5%	20.4%	
Owns a Car	86.7%	83.1%	0.207	89.9%	80.8%	0.015	88.3%	80.8%	.068
Has Both Computer and Internet at Home	84.3%	83.9%	0.900	85.2%	83.8%	0.728	83.9%	83.8%	.980
Ever Arrested	23.6%	17.8%	0.139	25.4%	16.3%	0.037	25.7%	16.3%	.034

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 test for significant imbalances for other variables.

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

Outcome	Full Sample	Survey Respondents	
		Unweighted Est (StdErr)	Weighted Est (StdErr)
Regular college credits earned during months:			
1-6	0.315	0.091 (0.405)	–0.041 (0.407)
7-12	0.382	0.195 (0.535)	–0.026 (0.529)
13-18	0.148	0.313 (0.530)	–0.042 (0.532)
1-18	0.844	0.600 (1.129)	–0.108 (1.144)
Proportion enrolled in college occupational training during months:			
1-6	0.0769**	0.0489 (0.0399)	0.0301 (0.0422)
7-12	0.0716**	0.0704* (0.0479)	0.0381 (0.0509)
13-18	0.0951**	0.1193*** (0.0495)	0.0881** (0.0521)
1-18	0.0653**	0.0391 (0.0359)	0.0153 (0.0387)
Total hours of occupational training at:			
A college	9.8	–3.4 (19.8)	–8.9 (19.6)
Proportion who received a credential from:			
A college	0.0076	–0.0239 (0.0398)	–0.0286 (0.0391)

SOURCE: Abt Associates calculations based on data from the administrative records and response status to the PACE short-term follow-up survey.

NOTES: Standard errors on estimated impacts are shown in parentheses. All estimates are regression adjusted, as discussed in Section B.3 of Appendix B. 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 10-percent level.

Appendix E: 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 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*. (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.)

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 et al. (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) and credits. The research team found no values that were clearly impossible, and thus retained all reported values in the analysis.