REPORT | February 2023

Exploring Employment and Education Outcomes for Caregivers Participating in Parents as Teachers

Home Visiting Outcomes Analysis Results

REPORT | February 2023

Exploring Employment and Education Outcomes for Caregivers Participating in Parents as Teachers

Home Visiting Outcomes Analysis Results

Authors

Yuan Wang, M.S., Nichole Sturmfels, M.P.H., James Bell Associates

Submitted to

Allison Kemner, M.P.H.
Senior Vice President and Chief Research Officer
Parents as Teachers National Center
2228 Ball Drive
St. Louis, MO 63146

Prepared by

James Bell Associates
2000 15th Street North, Suite 100
Arlington, VA 22201
(703) 528-3230
www.jbassoc.com

Nichole Sturmfels, Project Director

Contents

Exhibits

Overview

This study is a partnership between James Bell Associates (JBA) and Parents as Teachers (PAT) National Center to explore family economic well-being outcomes using secondary data. It examined outcomes in education, employment, and income for caregivers who participated in PAT relative to a comparison group from the Current Population Survey Annual Social and Economic Supplement (CPS-ASEC). This report describes the research questions and rationale, study design, analytic methods, and results from the multivariate analyses.

PAT model components support seven goals, including one focused on improving family economic well-being.[1] As such, goal setting with caregivers is a key intervention technique for the PAT model and may help drive family economic well-being. Analysis of caregiver goal data from PAT's data system, Penelope, revealed the top five goals set by caregivers at intake:[2] education, child development, basic needs, employment, and parenting behaviors.

We used a difference-in-difference study design to explore baseline and 1-year follow-up employment, education, and income measures over a 5-year pre–COVID-19 pandemic period for caregivers in both the PAT and comparison dataset. We then disaggregated the data by race and ethnicity and applied a moderation analysis to explore whether race and ethnicity moderated effects on employment and education outcomes.

Key Findings

One year after intake, caregivers in PAT* were—

- 16 percent more likely to become employed
- 69 percent more likely to enroll in high school
- 12 percent more likely to enroll in college

*Relative to the comparison group

This study included a racial equity impact analysis to examine and foster PAT's approach to serving marginalized families. It advances equity by examining racial and ethnic differences in employment and education outcomes for families receiving PAT home visiting. Given the historical impact of institutional and structural racism on opportunities, next steps should include gathering family and home visitor perspectives on the study results. Exploring potential theories of change could improve understanding of how home visiting interventions may help families overcome systemic barriers to achieving their employment and education goals.

---

[1] https://parentsasteachers.org/what-we-do/

[2] Within 90 days of enrollment.

Overall, we found statistically significant increases in employment and education for PAT caregivers relative to the comparison group. This suggests that caregivers participating in PAT are more likely to obtain employment and to enroll in high school or college relative to the comparison group. When we disaggregated the data by race and ethnicity, we found statistically significant differences in employment and education outcomes between and within racial and ethnic groups.

Methods

Study Design

We used a quasi-experimental design to evaluate change in employment, education, and income among PAT caregivers and a comparison group during a 1-year period. We further explored the differences in outcomes by disaggregating data by race and ethnicity.

The study was guided by the following research questions:

- Do employment, education, and income increase over time among PAT caregivers and a comparison group?
- Does race or ethnicity moderate the employment and education increase over time among PAT caregivers and a comparison group?
- Are some racial and ethnic subgroups affected more than others?

Sample

We extracted data from the PAT Penelope data system and the Current Population Survey Annual Social and Economic Supplement (CPS-ASEC) to construct the two groups for comparison. The PAT data contain mainly family-level data (e.g., home visits, income, characteristics) and individual caregiver data (e.g., demographics, screenings, health status, goals). The CPS-ASEC data contain information about family economic situations (e.g., educational attainment, poverty, income) and were obtained through the online portal Integrated Public Use Microdata Series.

This study analyzed data collected between January 1, 2015, and March 31, 2020, for both datasets. PAT requires caregiver employment and education and family income data to be collected at intake and updated as changes occur or at least annually thereafter. The CPS-ASEC data are reported once a year in March and include respondents' baseline and 1-year follow-up measures. Because both datasets contain baseline and follow-up data at approximately 1-year intervals, we were able to create a comparison group.

For both samples, data were included for participants aged 15–45 years, and only one participant from each family or household was included in the study (the primary caregiver in PAT or the head of the household in CPS-ASEC). Military respondents were excluded in the CPS-ASEC sample because they were not asked employment and income questions on the survey. The PAT sample was limited to affiliates actively using the Penelope data system for 12 months or more and excluded affiliates that use only the PAT curriculum and those participating in the Family and Child Education study because that study explored similar family economic well-being outcomes. To address family attrition bias, we excluded families who exited the program before the follow-up window (7 months after enrollment) started.

Measures

Predictors

Caregiver demographics included participant age, gender, race, ethnicity, marital status, education attainment, education and employment history, and income. All caregiver demographics were assessed at enrollment or baseline data collection. The treatment predictor (dichotomous) identifies data are from the PAT caregiver group. The post or posttreatment predictor (dichotomous) identifies that data are from the follow up time point.

Outcomes

The primary outcomes were employment status and enrollment in high school or college. The employment outcomes focused on those caregivers with both full-time and part-time employment, and the education outcomes focused on those caregivers who were enrolled in high school, a GED program, or college. High school enrollment was compared for those participants with an education attainment of less than a high school diploma or GED. College enrollment was compared for those who had a high school diploma or GED. A second outcome representing family income included family poverty status and ratio of income to poverty. Poverty status indicates if a family annual income is above or below the corresponding poverty line. The ratio of income to poverty was categorized as below 50 percent of poverty; 50 percent to 100 percent of poverty; 100 percent to 150 percent of poverty; 150 percent to 200 percent of poverty; and above 200 percent of poverty.

Covariates and Moderators

Race and ethnicity was self-identified and categorized as follows:

- Hispanic includes individuals of Hispanic ethnic origin of any race.
- Multiracial includes those of non-Hispanic ethnic origin who reported more than one race.

- Asian and Native Hawaiian or Other Pacific Islander (AAPCHO)[3] includes those of non-Hispanic ethnic origin who reported as Asian or Native Hawaiian or Other Pacific Islander.
- The remaining three groups are individuals of non-Hispanic ethnic origin who reported a single race: American Indian/Alaska Native (AI/AN), Black or African American, and White.

Analytic Strategy

We implemented a quasi-experimental difference-in-difference (DID) comparison between the PAT and comparison groups. Exploratory bivariate data analyses and a review of the literature informed the variables included in the initial DID model. We conducted separate analyses for each outcome of interest and used multiple imputation (MI) for missing data. We used SAS V.9.4 to perform the analyses.

To create a sound comparison, we used the entropy balancing technique to adjust the covariate distribution of the CPS-ASEC group data with a set of unit weights such that it becomes more like the covariate distribution in the PAT group, resulting in sample baseline equivalence. Entropy balancing provides superior covariate balance compared with traditional propensity score weighting methods because it does not require extensive iterative manual searching for a suitable weighting that balances the covariate distributions.

Next, we applied a binary logistic regression model and a proportional odds model with DID to compare the changes over time for employment, education, and income among PAT caregivers and CPS-ASEC participants. When randomization is not possible, DID can be used to imply causal inference. For example, we cannot determine causality by simply observing before-and-after changes in outcomes, because factors other than the treatment may influence the outcome over time. Further, we cannot simply compare enrolled and unenrolled groups because of selection bias and differences in unobservable characteristics between the groups. DID considers these limitations and allows for comparison of the before-and-after changes in outcomes for treatment and comparison groups to estimate the overall impact of the program. Finally, we conducted a moderation analysis to assess a differential effect of the treatment related to the different racial and ethnic subgroups.

[3] https://aapcho.org/

Data Analysis and Results

Univariate

We generated univariate descriptive statistics for each predictor and outcome variable to examine characteristics of participants in both groups (exhibit 1). After applying the inclusion criteria, 17,158 caregivers were in the PAT sample and 70,196 participants were in the CPS-ASEC sample. PAT caregivers were mostly female, young to middle age (72 percent aged 20–34 years), with limited income, and with low educational attainment (59 percent with a high school diploma, GED, or less). They were racially and ethnically diverse—a little more than a third of the caregivers in the sample were Hispanic or Latino. Among the rest of the sample, 40 percent were White, 20 percent were Black or African American, and 7 percent were other races. We reviewed measures of central tendency for continuous variables, age, and median income. On average, caregivers in the PAT sample were 29 years old (standard deviation (SD) = 6.5), and participants in the CPS-ASEC sample were 31 years old (SD = 8.7). The median family annual income[4] was $10,652 for PAT caregivers and $76,328 for CPS-ASEC participants.

Exhibit 1. Study participant characteristics

Frequency distribution	PAT		CPS-ASEC	
	%	*N*	%	*N*
Total		17,158		70,196
Age group				
15–19	7.4	1,263	11.5	9,393
20–24	20.3	3,475	15.7	9,035
25–29	27.0	4,640	17.9	10,509
30–34	24.5	4,199	17.5	12,353
35–45	20.9	3,581	37.4	28,906
Gender				
Male	14.7	2,527	50.2	34,150
Female	85.3	14,628	49.8	36,046

[4] Family annual income data are adjusted for inflation; for the PAT group, the outliers were detected and removed based on the interquartile range method.

Frequency distribution	PAT		CPS-ASEC	
	%	*N*	%	*N*
Race				
Black or African American	20.6	3,397	11.4	7117
White	68.0	11,198	78.9	55,832
American Indian/Alaska Native	4.2	695	1.1	1,011
Asian	2.2	366	6.2	4,355
Native Hawaiian or Other Pacific Islander	0.3	49	0.4	387
More than one race	4.6	755	2.0	1,494
Ethnicity				
Hispanic or Latino	33.4	5,649	15.7	12,082
Not Hispanic or Latino	66.6	11,275	84.3	58,114
Race and ethnicity				
Non-Hispanic, Black or African American	19.5	3,277	10.7	6,731
Non-Hispanic, White	40.4	6,820	65	44,754
Non-Hispanic, American Indian/Alaska Native	3.1	515	0.8	826
Non-Hispanic, Asian or Native Hawaiian or Other Pacific Islander	2.3	393	6.3	4,591
Non-Hispanic, Multiracial	1.4	228	1.6	1,212
Hispanic or Latino	33.4	5,649	15.7	12,082
Marital status				
Not married	51.7	7,623	58.8	37,956
Married	48.3	7,118	41.2	32,240
Employment status				
Not employed	49.9	7,576	28.2	20,157
Employed	50.1	7,579	71.8	50,039

Frequency distribution	PAT		CPS-ASEC	
	%	*N*	%	*N*
Education attainment				
Less than high school	25.0	3,820	2.3	1,763
High school diploma or GED	33.6	5,141	36.5	26,862
Some college or technical training	20.9	3,193	19.1	13,017
Associate's degree	4.1	633	9.2	6,533
Bachelor's degree or higher	16.3	2,493	32.9	22,021
High school enrollment				
Currently not enrolled	91.1	8,451	92.1	63,561
Currently enrolled	8.9	821	7.9	6,635
College enrollment				
Currently not enrolled	92.5	8,575	87.4	62,348
Currently enrolled	7.5	697	12.6	7,848
Poverty status				
Below poverty line	74.4	7,874	10.7	7,613
Above poverty line	24.8	2,620	89.3	62,583
Unknown	0.8	84	NA	NA
Ratio of income to poverty				
Below 50% of poverty	52.1	5,493	5.1	3,458
50% to 100% of poverty	22.4	2,364	5.5	4,127
100% to 150% of poverty	12.1	1,277	7.0	5,227
150% to 200% of poverty	5.9	622	7.8	5,764
200% and above of poverty	6.7	711	74.6	51,620
Unknown	0.8	84	NA	NA

Note: Comparison group number reports weighted proportions designated by the Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS). Missing data were excluded when calculating frequency. NA: Not Applicable or no data reported in the category

Bivariate

We used bivariate analyses (e.g., correlations, crosstabs) to understand data patterns and to refine our initial hypotheses and modeling strategy (exhibit 2).

Exhibit 2. Bivariate analytic strategy

Analysis	Purpose
Wilcoxon signed-rank tests	Explore change in the outcomes (employment, education, and family income)
Correlations, cross-tabulations, chi-square tests, logistic regressions	Explore relationships between predictors and the outcomes, including distribution and basic associations

First, we conducted a Wilcoxon signed-rank test to review whether outcomes change meaningfully over time from baseline to the follow-up assessment. No statistically significant difference existed (*P* = 0.23) between baseline and follow-up assessments for education attainment (exhibit 3). As a result, education attainment was excluded from the model where we examined education outcomes.

Exhibit 3. Wilcoxon signed-rank test results

Outcome domain	Outcome measurements	Signed-rank statistic (S)	*P* value
Employment	Employment status	50220	<.0001
Education	Education attainment	2056	**.2301**
Education	High school enrollment	–375	.0190
Education	College enrollment	38922	.0071
Family income	Family annual income	–4754000	<.0001
Family income	Poverty status	17605.5	<.0001
Family income	Ratio of income to poverty	83413.5	<.0001

Note: The Wilcoxon signed-rank test is a statistical hypothesis test used to test the location of a set of samples or to compare the locations of two populations using a set of matched samples. It is the nonparametric alternative test for the paired two-sample *t*-test.

Next, we reviewed how each predictor is related to each outcome. Correlations provide evidence of the size of associations between two variables, where correlation ≥|0.3| is considered modest and correlation ≥|0.8| is considered high. A predictor will be dropped in the multivariate model because of a lack of association with outcome. Exhibits 4 and 5 display the correlation results. Ethnicity had little relationship with high school enrollment and poverty status, and thus we dropped it from education and income analyses.

We then examined whether any predictors are highly correlated ($r \geq |0.8|$), which may indicate multicollinearity and can be problematic for regression analyses. As indicated in exhibit 4, the poverty status variable was highly correlated with the ratio of income to poverty variable ($r = 1$). We assessed multicollinearity using the variance inflation factor (VIF). The VIF estimates how much the variance of a regression coefficient is inflated due to multicollinearity. The general rule is that VIFs exceeding 10 indicate a sign of serious multicollinearity. Additionally, Akaike information criterion (AIC) or Bayesian information criterion (BIC) are two ways of scoring a model based on its log likelihood and complexity. Lower AIC or BIC values indicate a better-fit model, and a model with a 2-point decrease in AIC is considered significantly better than the model of comparison. If significant multicollinearity issues exist, we will retain the more robust variable that improves the fit of the model and will drop the other one from further analyses. After assessing the VIF and model fit using AIC and BIC, we retained the ratio of income to poverty variable and dropped poverty status variable from the employment and education analyses.

Exhibit 4. Spearman correlation tests and chi-square tests

Measurement 1	Measurement 2	Association test and statistic	Correlation coefficient	*P* value
Age	Family annual income	Spearman	0.16	<.0001
Age group	Sex	Chi-square, Cramer's V	0.16	<.0001
Age group	Race	Chi-square, Cramer's V	0.1	<.0001
Age group	Marital status	Chi-square, Cramer's V	0.37	<.0001
Age group	Ethnicity	Chi-square, Cramer's V	0.12	<.0001
Age group	Employment status	Chi-square, Cramer's V	0.11	<.0001
Age group	Education attainment	Chi-square, Cramer's V	0.19	<.0001
Age group	High school enrollment	Chi-square, Cramer's V	0.47	<.0001
Age group	College enrollment	Chi-square, Cramer's V	0.1	<.0001
Age group	Poverty status	Chi-square, Cramer's V	0.12	<.0001
Age group	Ratio of income to poverty	Chi-square, Cramer's V	0.1	<.0001
Sex	Race	Chi-square, Cramer's V	0.14	<.0001
Sex	Marital status	Chi-square, Cramer's V	–0.16	<.0001
Sex	Ethnicity	Chi-square, Cramer's V	–0.0003	.9654
Sex	Employment status	Chi-square, Cramer's V	0.33	<.0001

Measurement 1	Measurement 2	Association test and statistic	Correlation coefficient	*P* value
Sex	Education attainment	Chi-square, Cramer's V	0.12	<.0001
Sex	High school enrollment	Chi-square, Cramer's V	–0.1	<.0001
Sex	College enrollment	Chi-square, Cramer's V	–0.02	.1158
Sex	Poverty status	Chi-square, Cramer's V	0.1	<.0001
Sex	Ratio of income to poverty	Chi-square, Cramer's V	0.1	<.0001
Race	Marital status	Chi-square, Cramer's V	0.33	<.0001
Race	Ethnicity	Chi-square, Cramer's V	0.37	<.0001
Race	Employment status	Chi-square, Cramer's V	0.03	.0383
Race	Education attainment	Chi-square, Cramer's V	0.1	<.0001
Race	High school enrollment	Chi-square, Cramer's V	0.11	<.0001
Race	College enrollment	Chi-square, Cramer's V	0.1	<.0001
Race	Poverty status	Chi-square, Cramer's V	0.16	<.0001
Race	Ratio of income to poverty	Chi-square, Cramer's V	0.1	<.0001
Marital status	Ethnicity	Chi-square, Cramer's V	–0.13	<.0001
Marital status	Employment status	Chi-square, Cramer's V	0.05	<.0001
Marital status	Education attainment	Chi-square, Cramer's V	0.33	<.0001
Marital status	High school enrollment	Chi-square, Cramer's V	0.16	<.0001
Marital status	College enrollment	Chi-square, Cramer's V	0.03	.0253
Marital status	Poverty status	Chi-square, Cramer's V	0.23	<.0001
Marital status	Ratio of income to poverty	Chi-square, Cramer's V	0.25	<.0001
Ethnicity	Employment status	Chi-square, Cramer's V	–0.1	<.0001
Ethnicity	Education attainment	Chi-square, Cramer's V	0.29	<.0001
Ethnicity	High school enrollment	Chi-square, Cramer's V	–0.003	.7446
Ethnicity	College enrollment	Chi-square, Cramer's V	–0.1	<.0001
Ethnicity	Poverty status	Chi-square, Cramer's V	0.008	.4232

Measurement 1	Measurement 2	Association test and statistic	Correlation coefficient	*P* value
Ethnicity	Ratio of income to poverty	Chi-square, Cramer’s V	0.13	<.0001
Employment status	Education attainment	Chi-square, Cramer’s V	0.23	<.0001
Employment status	High school enrollment	Chi-square, Cramer’s V	–0.1	<.0001
Employment status	College enrollment	Chi-square, Cramer’s V	0.05	<.0001
Employment status	Poverty status	Chi-square, Cramer’s V	0.16	<.0001
Employment status	Ratio of income to poverty	Chi-square, Cramer’s V	0.19	<.0001
Education attainment	High school enrollment	Chi-square, Cramer’s V	0.26	<.0001
Education attainment	College enrollment	Chi-square, Cramer’s V	0.27	<.0001
Education attainment	Poverty status	Chi-square, Cramer’s V	0.26	<.0001
Education attainment	Ratio of income to poverty	Chi-square, Cramer’s V	0.17	<.0001
High school enrollment	College enrollment	Chi-square, Cramer’s V	–0.1	<.0001
High school enrollment	Poverty status	Chi-square, Cramer’s V	–0.1	<.0001
High school enrollment	Ratio of income to poverty	Chi-square, Cramer’s V	0.11	<.0001
College enrollment	Poverty status	Chi-square, Cramer’s V	0.05	.0001
College enrollment	Ratio of income to poverty	Chi-square, Cramer’s V	0.05	.0005
Poverty status	Ratio of income to poverty	Chi-square, Cramer’s V	1	<.0001

Exhibit 5. Logistic regression tests

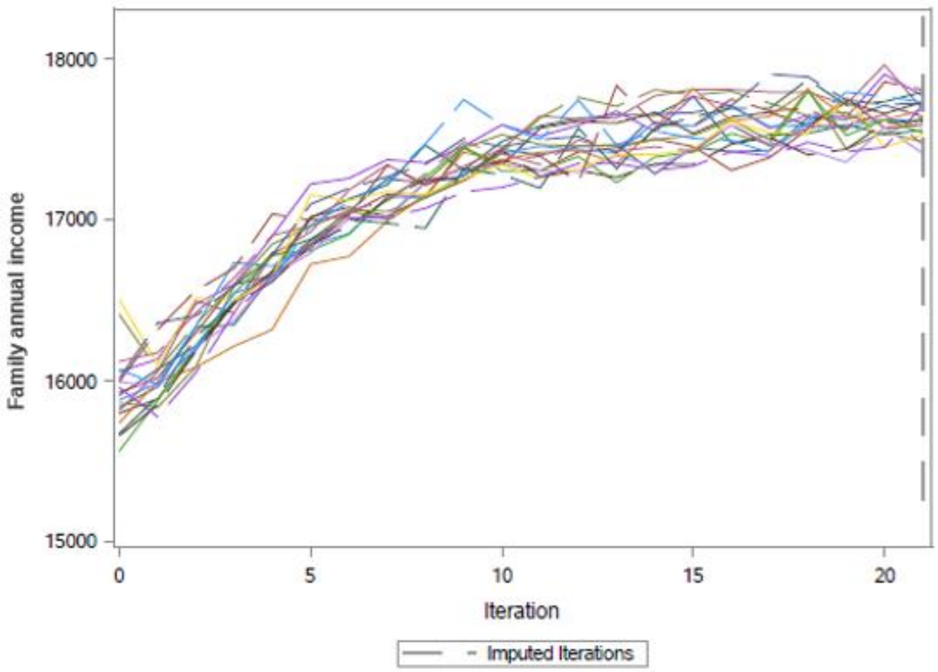
Dependent measurement	Independent measurement	Odds ratio	95% confidence interval	*P* value
Employment status	Age	1.024	[1.019, 1.029]	<.0001
Education attainment	Age	1.036	[1.031, 1.040]	<.0001
High school enrollment	Age	0.808	[0.795, 0.821]	<.0001
College enrollment	Age	0.976	[0.964, 0.988]	<.0001
Poverty status	Age	1.035	[1.028, 1.042]	<.0001
Ratio of income to poverty	Age	1.033	[1.027, 1.039]	<.0001

Missing Data

Upon evaluation of the missing data and patterns, we found 18 percent of data were missing at random and mostly occurring in the planned outcome variables. Because the data were missing at random, we addressed them using MI with the SAS Procedure (PROC) MI procedure. The following steps outline our process for MI.

Imputation or Fill-in Phase: Missing data points were imputed 20 times using a fully conditional specification approach (Liu & De, 2015). Fully conditional specification MI specifies the multivariate imputation model on a variable-by-variable basis and offers a principled yet flexible method of addressing missing data, which is particularly useful for large datasets with complex data structures. Categorical variables were imputed with the logistic regression method, and continuous variables were imputed with the predictive mean matching method. The quality of imputations was examined with graphic and numeric diagnostics (e.g., trace plots, Kolmogorov-Smirnov test) after the completion of imputations. Exhibit 6 displays the trace plot of means against the number of iterations for the family annual income variable. The plot examines the convergence and shows no apparent trends after 10 iterations. The dashed vertical lines indicate the imputed iterations.

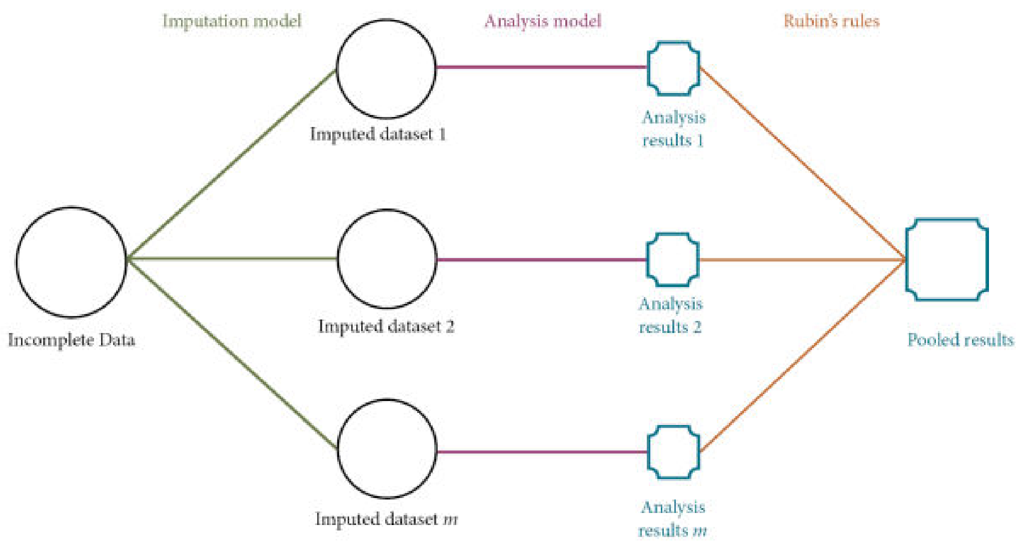
Exhibit 6. Trace plot of means for family annual income



Analysis Phase: All of the 20 "complete" datasets were then analyzed using (1) entropy balancing analysis to achieve sample baseline equivalence for covariates, (2) difference-in-difference analyses, and (3) difference-in-difference-in-difference analyses. Results for each dataset vary because of the difference in values assigned during the MI process.

Pooling Phase: The parameter estimates (e.g., coefficients, odds ratio, standard errors) obtained from each analyzed dataset are then combined to generate a single set of estimates for inference with Rubin's rules using the SAS PROC MIANALYZE procedure. Exhibit 7 depicts the three phases of MI.

Exhibit 7. Multiple imputation steps (m = 20)



Entropy Balancing

Randomized controlled trials are considered the gold standard for measuring the effectiveness of an intervention or treatment. This is because the act of randomization balances participant characteristics (both observed and unobserved) between the groups, thus allowing differences in the outcome to be attributed to the study intervention (Hainmueller, 2012). Because a randomized controlled trial was not feasible here, we used a quasi-experiental design with entropy balancing to design a sound comparison group.

Entropy balancing is a method used to adjust the covariate distribution of the comparison group data by reweighting the units such that it becomes similar to the covariate distribution in the treatment group. The purpose of using entropy balancing for this study is to diminish selection bias, remove differences in observed participants' characteristics between PAT and the comparison group, and reduce model dependency for the subsequent analysis of treatment effects.

Entropy balancing analysis was completed using the SAS PROC OPTMODEL procedure and the entropy balancing macro called *ebc* distibuted (Douglas et al., 2020). Exhibit 8 reports the nonentropy-weighted and entropy-weighted characteristics of the two groups at baseline. Employment, education, and income characteristics are missing in each respective column when they represent the outcome of interest. Proportion estimates with 95 percent confidence intervals were combined across 20 MI datasets.

Exhibit 8. Pooled characteristics of PAT and comparison caregivers, nonentropy weighted vs. entropy weighted

Pooled participant characteristics	Nonentropy weighted		Entropy weighted					
			Employment outcomes		Education outcomes		Family income outcomes	
	PAT	Comparison	PAT	Comparison	PAT	Comparison	PAT	Comparison
	%	%	%	%	%	%	%	%
	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Age group (%)								
15–19	7.4	11.5	7.4	7.4	7.4	7.4	7.4	7.4
20–24	20.3	15.7	20.3	20.3	20.3	20.3	20.3	20.3
25–29	27	17.9	27	27	27	27	27	27
30–34	24.5	17.5	24.5	24.5	24.5	24.5	24.5	24.5
35–45	20.9	37.4	20.9	20.9	20.9	20.9	20.9	20.9
Sex (%)								
Male	14.7 (14.2, 15.3)	50.2	14.7 (14.2, 15.3)	14.7 (13.6, 15.9)	14.7 (14.2, 15.3)	14.7 (13.5, 16.0)	14.7 (14.2, 15.3)	14.7 (14.1, 15.3)
Female	85.3 (84.7, 85.8)	49.8	85.3 (84.7, 85.8)	85.3 (84.1, 86.4)	85.3 (84.7, 85.8)	85.3 (84.0, 86.5)	85.3 (84.7, 85.8)	85.3 (84.7, 85.9)
Race (%)								
Black or African American	20.2 (19.5, 20.8)	11.4	20.2 (19.5, 20.8)	20.2 (18.3, 22)	20.2 (19.5, 20.8)	20.2 (18.4, 21.9)	20.2 (19.5, 20.8)	20.2 (18.2, 22.2)
White	68.3 (67.6, 69.0)	78.9	68.3 (67.6, 69.0)	68.3 (66.1, 70.6)	68.3 (67.6, 69.0)	68.3 (66.1, 70.6)	68.3 (67.6, 69.0)	68.3 (66.2, 70.5)
American Indian/Alaska Native	4.2 (3.9, 4.5)	1.1	4.2 (3.9, 4.5)	4.2 (3, 5.3)	4.2 (3.9, 4.5)	4.2 (3.2, 5.1)	4.2 (3.9, 4.5)	4.2 (3.0, 5.4)

Pooled participant characteristics	Nonentropy weighted		Entropy weighted					
			Employment outcomes		Education outcomes		Family income outcomes	
	PAT	Comparison	PAT	Comparison	PAT	Comparison	PAT	Comparison
	%	%	%	%	%	%	%	%
	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Asian	2.2 (2.0, 2.4)	6.2	2.2 (2.0, 2.4)	2.2 (1.7, 2.7)	2.2 (2.0, 2.4)	2.2 (1.6, 2.7)	2.2 (2.0, 2.4)	2.2 (2.0, 2.4)
Native Hawaiian or Other Pacific Islander	0.3 (0.2, 0.4)	0.4	0.3 (0.2, 0.4)	0.3 (0.2, 0.4)	0.3 (0.2, 0.4)	0.3 (0.2, 0.4)	0.3 (0.2, 0.4)	0.3 (0.2, 0.4)
More than one race	4.9 (4.5, 5.2)	2	4.9 (4.5, 5.2)	4.9 (3.5, 6.2)	4.9 (4.5, 5.2)	4.9 (3.5, 6.2)	4.9 (4.5, 5.2)	4.9 (3.5, 6.2)
Ethnicity (%)								
Not Hispanic or Latino	66.7 (66.0, 67.4)	84.3	66.7 (66.0, 67.4)	66.7 (64.1, 69.3)	66.7 (66.0, 67.4)	66.7 (64.0, 69.4)	66.7 (66.0, 67.4)	66.7 (65.0, 68.5)
Hispanic or Latino	33.3 (32.6, 34.0)	15.7	33.3 (32.6, 34.0)	33.3 (30.7, 35.9)	33.3 (32.6, 34.0)	33.3 (30.6, 36.0)	33.3 (32.6, 34.0)	33.3 (31.5, 35.0)
Race and ethnicity (%)								
Non-Hispanic, Black or African American	19.5 (18.9, 20.1)	10.7	19.5 (18.9, 20.1)	19.5 (17.8, 21.3)	19.5 (18.9, 20.1)	19.5 (17.8, 21.2)	19.5 (18.9, 20.1)	19.5 (17.5, 21.5)
Non-Hispanic, White	40.4 (39.6, 41.1)	65	40.4 (39.6, 41.1)	40.4 (38.0, 42.7)	40.4 (39.6, 41.1)	40.4 (37.9, 42.8)	40.4 (39.6, 41.1)	40.4 (38.8, 41.9)
Non-Hispanic, American Indian/Alaska Native	3.1 (2.9, 3.4)	0.8	3.1 (2.9, 3.4)	3.1 (2.4, 3.8)	3.1 (2.9, 3.4)	3.1 (2.4, 3.9)	3.1 (2.9, 3.4)	3.1 (2.2, 4.0)
Non-Hispanic, Asian or Native Hawaiian or Other Pacific Islander	2.4 (2.1, 2.6)	6.3	2.4 (2.1, 2.6)	2.4 (1.9, 2.8)	2.4 (2.1, 2.6)	2.4 (1.8, 2.9)	2.4 (2.1, 2.6)	2.4 (2.1, 2.6)
Non-Hispanic, Multiracial	1.4 (1.2, 1.5)	1.6	1.4 (1.2, 1.5)	1.4 (1.1, 1.7)	1.4 (1.2, 1.5)	1.4 (1.0, 1.7)	1.4 (1.2, 1.5)	1.4 (1.1, 1.6)

Pooled participant characteristics	Nonentropy weighted		Entropy weighted					
			Employment outcomes		Education outcomes		Family income outcomes	
	PAT	Comparison	PAT	Comparison	PAT	Comparison	PAT	Comparison
	%	%	%	%	%	%	%	%
	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Hispanic or Latino	33.3 (32.6, 34.0)	15.7	33.3 (32.6, 34.0)	33.3 (30.7, 35.9)	33.3 (32.6, 34.0)	33.3 (30.6, 36.0)	33.3 (32.6, 34.0)	33.3 (31.5, 35.0)
Marital status (%)								
Not married	51.7 (50.9, 52.5)	58.8	51.7 (50.9, 52.5)	51.7 (49.2, 54.1)	51.7 (50.9, 52.5)	51.7 (49.1, 54.3)	51.7 (50.9, 52.5)	51.7 (49.9, 53.4)
Married	48.3 (47.5, 49.1)	41.2	48.3 (47.5, 49.1)	48.3 (45.9, 50.8)	48.3 (47.5, 49.1)	48.3 (45.7, 50.9)	48.3 (47.5, 49.1)	48.3 (46.6, 50.1)
Employment status (%)								
Not employed	50.4 (49.6, 51.2)	28.2			50.4 (49.6, 51.2)	50.4 (47.8, 53.0)	50.4 (49.6, 51.2)	50.4 (48.6, 52.2)
Employed	49.6 (48.8, 50.4)	71.8			49.6 (48.8, 50.4)	49.6 (47.0, 52.2)	49.6 (48.8, 50.4)	49.6 (47.8, 51.4)
Education attainment (%)								
Less than high school	25.2 (24.3, 26.0)	2.3	25.2 (24.3, 26.0)	25.2 (22.1, 28.2)	25.2 (24.3, 26.0)	25.2 (21.9, 28.5)	25.2 (24.3, 26.0)	25.2 (22.6, 27.8)
High school diploma or GED	33.7 (32.9, 34.5)	36.5	33.7 (32.9, 34.5)	33.7 (31.7, 35.7)	33.7 (32.9, 34.5)	33.7 (31.6, 35.8)	33.7 (32.9, 34.5)	33.7 (32.3, 35.1)
Some college or technical training	20.9 (20.2, 21.6)	19.1	20.9 (19.3, 22.5)	20.9 (20.2, 21.6)	20.9 (19.3, 22.5)	20.9 (20.2, 21.6)	20.9 (19.9, 21.9)	20.9 (20.2, 21.6)
Associate's degree	4.1 (3.6, 4.6)	9.2	4.1 (3.6, 4.6)	4.1 (3.5, 4.8)	4.1 (3.6, 4.6)	4.1 (3.5, 4.7)	4.1 (3.6, 4.6)	4.1 (3.7, 4.6)
Bachelor's degree or higher	16.1 (15.4, 16.9)	32.9	16.1 (15.4, 16.9)	16.1 (14.7, 17.6)	16.1 (15.4, 16.9)	16.1 (14.7, 17.6)	16.1 (15.4, 16.9)	16.1 (15.3, 16.9)

Pooled participant characteristics	Nonentropy weighted		Entropy weighted					
			Employment outcomes		Education outcomes		Family income outcomes	
	PAT	Comparison	PAT	Comparison	PAT	Comparison	PAT	Comparison
	%	%	%	%	%	%	%	%
	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
High school enrollment (%)								
Currently not enrolled	92 (91.6, 92.5)	92.1	92 (91.6, 92.5)	92 (90.4, 93.7)			92 (91.6, 92.5)	92 (90.5, 93.6)
Currently enrolled	8 (7.5, 8.4)	7.9	8 (7.5, 8.4)	8 (6.3, 9.6)			8 (7.5, 8.4)	8 (6.4, 9.5)
College enrollment (%)								
Currently not enrolled	92.2 (91.6, 92.8)	87.4	92.2 (91.6, 92.8)	92.2 (91.3, 93)			92.2 (91.6, 92.8)	92.2 (91.6, 92.8)
Currently enrolled	7.8 (7.2, 8.4)	12.6	7.8 (7.2, 8.4)	7.8 (7, 8.7)			7.8 (7.2, 8.4)	7.8 (7.2, 8.4)
Poverty status (%)								
Below poverty line	72.8 (72.0, 73.6)	10.9	72.8 (72, 73.6)	72.8 (71.3, 74.3)	72.8 (72.0, 73.6)	72.8 (71.2, 74.3)		
Above poverty line	27.2 (26.4, 28.0)	89.1	27.2 (26.4, 28)	27.2 (25.7, 28.7)	27.2 (26.4, 28.0)	27.2 (25.7, 28.8)		
Ratio of income to poverty (%)								
Below 50% of poverty	51.4 (50.4, 52.3)	5.3	51.4 (50.4, 52.3)	51.4 (48.9, 53.8)	51.4 (50.4, 52.3)	51.4 (48.8, 53.9)		
50% to 100% of poverty	21.4 (20.7, 22.2)	5.7	21.4 (20.7, 22.2)	21.4 (19.7, 23.2)	21.4 (20.7, 22.2)	21.4 (19.8, 23.1)		
100% to 150% of poverty	12.4 (11.8, 13.1)	7.1	12.4 (11.8, 13.1)	12.4 (11.5, 13.4)	12.4 (11.8, 13.1)	12.4 (11.5, 13.4)		

Pooled participant characteristics	Nonentropy weighted		Entropy weighted					
			Employment outcomes		Education outcomes		Family income outcomes	
	PAT	Comparison	PAT	Comparison	PAT	Comparison	PAT	Comparison
	%	%	%	%	%	%	%	%
	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
150% to 200% of poverty	6.5 (6.1, 7.0)	7.9	6.5 (6.1, 7.0)	6.5 (6, 7)	6.5 (6.1, 7.0)	6.5 (6.1, 7.0)		
200% and above of poverty	8.2 (7.7, 8.8)	74	8.2 (7.7, 8.8)	8.2 (7.6, 8.8)	8.2 (7.7, 8.8)	8.2 (7.6, 8.8)		
Income (mean)								
Family annual income	18,180 (17,607; 18,753)	98,623	18,180 (17,607; 18,753)	18,180 (17,296; 19,064)	18,180 (17,607; 18,753)	18,180 (17,290; 19,070)		

Note: Nonentropy-weighted participants in the comparison group report weighted proportions designated by the Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS). Ninety-five percent confidence intervals are displayed below the percentages and mean to account for the uncertainty caused by the missing data.

Multivariate

Difference-in-Difference Analysis

To answer our first research question, we conducted logistic regression with a DID comparison between the PAT and comparison groups for each outcome of interest. This approach controls for unobservable time and group characteristics that confound the effect of the treatment on the outcome.

- Do employment, education, and income increase over time among PAT caregivers and a comparison group?

Key Concept: DID is a quasi-experimental statistical technique used to estimate treatment effects by comparing the change in the differences in observed outcomes between treatment and comparison groups, across pretreatment and posttreatment periods. The core theory of DID is that the difference between treatment and control groups in the change from pretreatment to posttreatment can be interpreted as the effect of treatment.

Calculation: The general structure of the DID is presented in Exhibit 9. DID can be calculated as the difference between the two groups in the changes in time of the group means. More clearly, the DID estimator takes the difference in the treatment group before and after the treatment ($\bar{y}_{T,Post} - \bar{y}_{T,Pre}$) and subtracts the difference in the control group before and after the treatment ($\bar{y}_{C,Post} - \bar{y}_{C,Pre}$), as in the following formula: $DID = (\bar{y}_{T,Post} - \bar{y}_{T,Pre}) - (\bar{y}_{C,Post} - \bar{y}_{C,Pre})$.

Exhibit 9. Difference-in-difference structure

	Pretreatment	Posttreatment	Change from before to after treatment
Treatment group	$\bar{y}_{T,Pre}$	$\bar{y}_{T,Post}$	$\bar{y}_{T,Post} - \bar{y}_{T,Pre}$
Control group	$\bar{y}_{C,Pre}$	$\bar{y}_{C,Post}$	$\bar{y}_{C,Post} - \bar{y}_{C,Pre}$
Difference within period	$\bar{y}_{T,Pre} - \bar{y}_{C,Pre}$	$\bar{y}_{T,Post} - \bar{y}_{C,Post}$	$(\bar{y}_{T,Post} - \bar{y}_{T,Pre}) - (\bar{y}_{C,Post} - \bar{y}_{C,Pre})$

Regression Model: Impacts calculated based on DID are usually derived within a regression framework that also accounts for other observed covariates. The regression model–implemented DID design can be defined as

$$Y = \beta_0 + \beta_1 Time + \beta_2 Treatment + \boldsymbol{\beta_3 Time} \times \boldsymbol{Treatment} + \beta_4 Covaritate + \varepsilon$$

where Y is the outcome variable, $Treatment$ is a dummy variable indicating the treatment and comparison group, and $Time$ is a dummy variable indicating pre- and posttreatment. The coefficients can be interpreted as follows:

β_0: Average outcome of the control group before the treatment ($\bar{y}_{C,Pre}$)

β_1: Average change of control group over time $(\bar{y}_{C,Post} - \bar{y}_{C,Pre})$

β_2: Difference between the two groups before the treatment $(\bar{y}_{T,Pre} - \bar{y}_{C,Pre})$

β_3: Difference in changes between the two groups over time $(\bar{y}_{T,Post} - \bar{y}_{T,Pre}) - (\bar{y}_{C,Post} - \bar{y}_{C,Pre})$

Model Building

A binary logistic regression model or a proportional odds model was fit to compare outcomes in employment, education, and income among PAT caregivers and comparison participants. The entropy-weighted characteristics were included as covariates in the DID models. Analyses were performed using the SAS PROC LOGISTIC procedure.

Employment Status. To explore PAT's impact on employment status over time, we performed logistic regression analysis with the DID design. Exhibit 10 displays the models explored. Exhibit 11 displays the best-fitting and most parsimonious model for the employment status outcome.

Exhibit 10. Model comparison for employment status

DID estimates for employment status			
Model	**Odds ratio**	**95% confidence interval**	***P* value**
Basic DID model	0.803	(0.765, 0.842)	<.0001
Adjusted DID model	1.164	(1.104, 1.228)	<.0001
Parsimonious adjusted DID model	1.164	(1.104, 1.228)	<.0001

Note: The basic DID model includes only treatment, post, and the interaction between post and treatment as predictors. The adjusted DID model includes treatment, post, interaction (treatment x post), and all other covariates. The parsimonious adjusted DID model includes treatment, post, interaction (treatment x post), and all other statistically significant covariates ($P < .05$).

Following is the equation of the final DID model. Exhibit 11 reports the pooled estimates for the employment outcome and provides odds ratios and the 95% confidence interval.

$$Logit\ \Pr(Employed) = \beta_0 + \beta_1 Time + \beta_2 Treatment + \beta_3 Time \times Treatment + \beta_4 Age\ Group + \beta_5 Gender + \beta_6 Race + \beta_7 Marital\ Status + \beta_8 Education\ Attainment + \beta_9 High\ School\ Enrollment + \beta_{10} College\ Enrollment + \beta_{11} Ratio\ of\ Income\ to\ Poverty + \beta_{12} Age\ Group \times High\ School\ Enrollment + \varepsilon$$

PAT caregivers who were not employed at baseline were 16 percent more likely to be employed after 1 year than those in the comparison group. In other words, PAT caregivers had a statistically significant increase in employment over time (odds ratio (OR) = 1.16, 95% confidence interval (CI) = 1.104 to 1.228, $p < .05$) relative to individuals in the comparison group.

Exhibit 11. Pooled estimates for employment status

Employment status (modeled on employed)	Odds ratio	95% confidence interval	*P* value statistical significance
post x treatment (DID estimates)	1.164	(1.104, 1.228)	**
treatment	1.663	(1.594, 1.735)	**
post	0.89	(0.868, 0.913)	**
Age group (reference: 15–19)			
20–24	1.569	(1.460, 1.687)	**
25–29	1.392	(1.298, 1.493)	**
30–34	1.602	(1.478, 1.736)	**
35–45	1.47	(1.363, 1.586)	**
Gender (reference: male)			
Female	0.295	(0.286, 0.305)	**
Race (reference: Black or African American)			
White	0.880	(0.853, 0.907)	**
American Indian/Alaska Native	1.022	(0.954, 1.096)	
Asian	0.852	(0.789, 0.921)	**
Native Hawaiian or Other Pacific Islander	0.771	(0.628, 0.948)	**
More than one race	1.033	(0.973, 1.097)	
Marital status (reference: not married)			
Married	0.680	(0.657, 0.704)	**
Education attainment (reference: less than high school)			
High school diploma or GED	1.052	(1.015, 1.091)	**
Some college or technical training	1.450	(1.394, 1.507)	**
Associate’s degree	1.584	(1.495, 1.679)	**
Bachelor’s degree or higher	1.713	(1.641, 1.787)	**
High school enrollment (reference: currently not enrolled)			
Currently enrolled	0.286	(0.258, 0.316)	**
College enrollment (reference: currently not enrolled)			
Currently enrolled	0.672	(0.640, 0.705)	**
Ratio of income to poverty (reference: below 50% of poverty)			
50% to 100% of poverty	2.113	(2.031, 2.200)	**
100% to 150% of poverty	3.124	(3.012, 3.240)	**
150% to 200% of poverty	3.908	(3.736, 4.089)	**
200% and above poverty	5.581	(5.340, 5.832)	**
Age group and high school enrollment (reference: 15–19 and currently not enrolled)			
20–24 and enrolled	0.587	(0.476, 0.725)	**
25–29 and enrolled	0.881	(0.696, 1.114)	

Employment status (modeled on employed)	Odds ratio	95% confidence interval	*P* value statistical significance
30–34 and enrolled	13.822	(10.995, 17.375)	**
35–45 and enrolled	1.732	(1.304, 2.301)	**

Note: **Statistical significance at $p < .05$.

High School Enrollment. To explore PAT's impact on high school enrollment over time, we performed logistic regression analysis with the DID design. Exhibit 12 displays the models explored. Exhibit 13 displays the best-fitting and most parsimonious model for high school enrollment outcome.

Exhibit 12. Model comparison for high school enrollment

DID estimates for high school enrollment			
Model	**Odds ratio**	**95% confidence interval**	***P* value**
Basic DID model	1.315	(1.198, 1.443)	<.0001
Full DID model	1.692	(1.475, 1.941)	<.0001
Parsimonious DID model	1.690	(1.474, 1.938)	<.0001

Note: The basic DID model includes only treatment, post, and the interaction between post and treatment as predictors. The full DID model includes treatment, post, interaction (treatment x post), and all other covariates. The parsimonious DID model includes treatment, post, interaction (treatment x post), and all other statistically significant covariates ($P < .05$).

Following is the equation of the final DID model. Exhibit 13 reports the pooled estimates for the high school enrollment outcome and provides odds ratios and 95 percent confidence intervals.

$$Logit\ \Pr(High\ School\ Enrolled) = \beta_0 + \beta_1 Time + \beta_2 Treatment + \beta_3 Time \times Treatment + \beta_4 Age\ Group + \beta_5 Race + \beta_6 Marital\ Status + \beta_7 Education\ Attainment + \beta_8 Employment\ Status + \beta_9 Ratio\ of\ Income\ to\ Poverty + \beta_{10} Marital\ Status \times Eduation\ Attainment + \varepsilon$$

PAT caregivers who were not enrolled in high school at baseline were 69 percent more likely to be enrolled in high school after 1 year than those in the comparison group. In other words, PAT caregivers had a statistically significant increase in high school enrollment over time (OR = 1.69, 95% CI = 1.474 to 1.938, $p < .05$) relative to individuals in the comparison group.

Exhibit 13. Pooled estimates for high school enrollment

High school enrollment (modeled on high school enrolled)	Odds ratio	95% confidence interval	*P* value statistical significance
post x treatment (DID estimates)	1.690	(1.474, 1.938)	**
treatment	1.778	(1.594, 1.983)	**
post	0.584	(0.541, 0.631)	**

High school enrollment (modeled on high school enrolled)	Odds ratio	95% confidence interval	*P* value statistical significance
Age group (reference: 15–19)			
20–24	0.042	(0.037, 0.046)	**
25–29	0.011	(0.010, 0.012)	**
30–34	0.017	(0.015, 0.019)	**
35–45	0.010	(0.008, 0.012)	**
Race (reference: Black or African American)			
White	0.602	(0.552, 0.656)	**
American Indian/Alaska Native	0.738	(0.633, 0.859)	**
Asian	0.674	(0.497, 0.915)	**
Native Hawaiian or Other Pacific Islander	0.592	(0.322, 1.088)	*
More than one race	0.699	(0.590, 0.827)	**
Marital status (reference: not married)			
Married	0.200	(0.161, 0.249)	**
Education attainment (reference: less than high school)			
High school diploma or GED	0.735	(0.673, 0.803)	**
Some college or technical training	0.073	(0.061, 0.086)	**
Associate's degree	0.057	(0.023, 0.140)	**
Bachelor's degree or higher	0.068	(0.038, 0.122)	**
Employment status (reference: not employed)			
Employed	0.405	(0.375, 0.437)	**
Ratio of income to poverty (reference: below 50% of poverty)			
50% to 100% of poverty	1.199	(1.099, 1.309)	**
100% to 150% of poverty	0.851	(0.745, 0.973)	**
150% to 200% of poverty	0.986	(0.834, 1.165)	
200% and above poverty	0.977	(0.849, 1.125)	
Marital status and education attainment (reference: not married and less than high school)			
Married and high school diploma or GED	0.97	(0.746, 1.262)	
Married and some college or technical training	3.493	(1.760, 6.933)	**
Married with associate’s degree	6.121	(2.114, 17.724)	**
Married with bachelor’s degree or higher	4.619	(1.877, 11.371)	**

Note: *Statistical significance at $p < .1$. **Statistical significance at $p < .05$.

College Enrollment. To explore the PAT intervention's impact on college enrollment over time, we performed logistic regression analysis with the DID design. Exhibit 14 displays the models explored. Exhibit 15 displays the best-fitting and most parsimonious model for college enrollment.

Exhibit 14. Model comparison for college enrollment

DID estimates for college enrollment			
Model	Odds ratio	95% confidence interval	*P* value
Basic DID model	1.099	(0.988, 1.222)	.081
Full DID model	1.106	(0.986, 1.241)	.086
Parsimonious DID model	1.118	(0.995, 1.256)	.061

Note: The basic DID model includes only treatment, post, and the interaction between post and treatment as predictors. The full DID model includes treatment, post, interaction (treatment x post), and all other covariates. The parsimonious DID model includes treatment, post, interaction (treatment x post), and all other statistically significant covariates ($P < .05$).

Following is the equation of the final DID model. Exhibit 15 reports the pooled estimates for the college enrollment outcome and provides odds ratios and the 95% confidence interval.

$$Logit\ \Pr(College\ Enrolled) = \beta_0 + \beta_1 Time + \beta_2 Treatment + \beta_3 Time \times Treatment + \beta_4 Age\ Group + \beta_5 Gender + \beta_6 Race + \beta_7 Marital\ Status + \beta_8 Education\ Attainment + \beta_9 Employment\ Status + \beta_{10} Gender\ \times\ Employment\ Status + \varepsilon$$

PAT caregivers who were not enrolled in college at baseline were 12 percent more likely to be enrolled in college after 1 year than those in the comparison group. In other words, PAT caregivers had a statistically significant increase in enrolling in college over time (OR = 1.12, 95% CI = 0.995 to 1.256, $p < .1$) relative to those individuals in the comparison group.

Exhibit 15. Pooled estimates for college enrollment

College enrollment (modeled on college enrolled)	Odds ratio	95% confidence interval	*P* value statistical significance
post x treatment (DID estimates)	1.118	(0.995, 1.256)	*
treatment	0.779	(0.707, 0.858)	**
post	0.903	(0.870, 0.938)	**
Age group (reference: 15–19)			
20–24	0.443	(0.416, 0.471)	**
25–29	0.171	(0.160, 0.184)	**
30–34	0.124	(0.115, 0.133)	**
35–45	0.070	(0.064, 0.077)	**
Gender (reference: male)			
Female	0.821	(0.738, 0.913)	**

College enrollment (modeled on college enrolled)	Odds ratio	95% confidence interval	*P* value statistical significance
Race (reference: Black or African American)			
White	1.216	(1.163, 1.270)	**
American Indian/Alaska Native	1.068	(0.972, 1.174)	
Asian	1.877	(1.678, 2.100)	**
Native Hawaiian or Other Pacific Islander	0.410	(0.212, 0.793)	**
More than one race	1.547	(1.405, 1.704)	**
Marital status (reference: not married)			
Married	0.478	(0.456, 0.501)	**
Education attainment (reference: less than high school)			
High school diploma or GED	1.376	(1.275, 1.486)	**
Some college or technical training	13.073	(12.142, 14.076)	**
Associate's degree	8.882	(7.902, 9.984)	**
Bachelor's degree or higher	11.324	(10.412, 12.315)	**
Employment status (reference: not employed)			
Employed	0.431	(0.381, 0.487)	**
Gender and employment status (reference: male and not employed)			
Female and employed	1.508	(1.328, 1.713)	**

Note: *Statistical significance at $p < .1$. **Statistical significance at $p < .05$.

Poverty Status. To explore PAT's impact on poverty status over time, we performed logistic regression analysis with the DID design. Exhibit 16 displays the models explored. Exhibit 17 displays the best-fitting and most parsimonious model (in this case, it is the full model) for poverty status.

Exhibit 16. Model comparison for poverty status

DID estimates for poverty status			
Model	Odds ratio	95% confidence interval	*P* value
Basic DID model	0.970	(0.915, 1.028)	.305
Full DID model	0.999	(0.935, 1.069)	.985

Note: The basic DID model includes only treatment, post, and the interaction between post and treatment as predictors. The full DID model includes treatment, post, interaction (treatment x post), and all other covariates.

Following is the equation of the final DID model. Exhibit 17 reports the pooled estimates for poverty status outcome and provides odds ratios and the 95% confidence interval.

$$Logit\ \Pr(Above\ Poverty) = \beta_0 + \beta_1 Time + \beta_2 Treatment + \beta_3 Time \times Treatment + \beta_4 Age\ Group + \beta_5 Race + \beta_6 Ethnicity + \beta_7 Marital\ Status + \beta_8 Education\ Attainment + \beta_9 Employment\ Status + \beta_{10} High\ School\ Enrollment + \beta_{11} College\ Enrollment + \varepsilon$$

The effect of PAT on poverty status over time was not statistically significant (OR = 0.99, 95% CI = 0.935 to 1.069) compared to the comparison group.

Exhibit 17. Pooled estimates for poverty status

Poverty status (modeled on above poverty)	Odds ratio	95% confidence interval	*P* value statistical significance
post x treatment (DID estimates)	0.999	(0.935, 1.069)	
treatment	0.059	(0.055, 0.063)	**
post	1.088	(1.052, 1.125)	**
Age group (reference: 15–19)			
20–24	1.096	(0.974, 1.234)	
25–29	0.629	(0.560, 0.705)	**
30–34	0.488	(0.451, 0.529)	**
35–45	0.651	(0.600, 0.706)	**
Gender (reference: male)			
Female	0.684	(0.648, 0.723)	**
Race (reference: Black or African American)			
White	1.752	(1.671, 1.836)	**
American Indian/Alaska Native	1.014	(0.877, 1.172)	
Asian	1.883	(1.648, 2.151)	**
Native Hawaiian or Other Pacific Islander	1.235	(0.884, 1.725)	
More than one race	1.173	(0.860, 1.600)	
Ethnicity (reference: not Hispanic or Latino)			
Hispanic or Latino	0.735	(0.667, 0.809)	**
Marital status (reference: not married)			
Married	2.60	(2.502, 2.703)	**
Education attainment (reference: less than high school)			
High school diploma or GED	1.526	(1.375, 1.695)	**
Some college or technical training	2.672	(2.408, 2.965)	**
Associate's degree	3.607	(3.290, 3.954)	**
Bachelor's degree or higher	6.752	(6.117, 7.453)	**
Employment status (reference: not employed)			
Employed	2.976	(2.877, 3.078)	**
High school enrollment (reference: currently not enrolled)			
Currently enrolled	2.334	(2.169, 2.511)	**
College enrollment (reference: currently not enrolled)			
Currently enrolled	1.126	(1.036, 1.223)	**

Note: **Statistical significance at $p < 0.05$.

Ratio of Income to Poverty. To measure PAT's impact on the ratio of income to poverty over time, we fit a proportional odds model with the DID design. Exhibit 18 displays the models explored. Exhibit 19 displays the best-fitting and most parsimonious model (in this case, it is the full model).

Exhibit 18. Model comparison for ratio of income to poverty

DID estimates for ratio of income to poverty			
Model	Odds ratio	95% confidence interval	*P* value
Basic DID model	1.014	(0.960, 1.070)	.619
Full DID model	1.043	(0.982, 1.108)	.174

Note: The basic DID model includes only treatment, post, and the interaction between post and treatment as predictors. The full DID model includes treatment, post, interaction (treatment x post), and all other covariates.

Following is the equation of the final DID model. Exhibit 19 reports the pooled estimates for ratio of income to poverty outcome and provides odds ratios and the 95% confidence interval.

$$Logit\ \Pr(Ratio\ of\ Income\ to\ Poverty\ \leq j) = \beta_{j0} + \beta_{j1}Time + \beta_{j2}Treatment + \beta_{j3}Time \times Treatment + \beta_{j4}Age\ Group + \beta_{j5}Race + \beta_{j6}Ethnicity + \beta_{j7}Marital\ Status + \beta_{j8}Education\ Attainment + \beta_{j9}Employment\ Status + \beta_{j10}High\ School\ Enrollment + \beta_{j11}College\ Enrollment + \varepsilon_j$$

$$where\ j = < 50\%, \quad 50\% - 100\%, \quad 100\% - 150\%, \quad 150\% - 200\%$$

The effect of PAT on the ratio of income to poverty over time was not statistically significant (OR = 1.04, 95% CI = 0.982 to 1.108) compared to the control group.

Exhibit 19. Pooled estimates for ratio of income to poverty

Ratio of income to poverty (modeled on higher ratio)	Odds ratio	95% confidence interval	*P* value statistical significance
post x treatment (DID estimates)	1.043	(0.982, 1.108)	
treatment	0.068	(0.064, 0.073)	**
post	1.035	(0.990, 1.081)	
Age group (reference: 15–19)			
20–24	0.990	(0.856, 1.145)	
25–29	0.607	(0.555, 0.665)	**
30–34	0.517	(0.484, 0.552)	**
35–45	0.660	(0.611, 0.713)	**
Gender (reference: male)			
Female	0.700	(0.667, 0.735)	**
Race (reference: Black or African American)			
White	1.830	(1.720, 1.946)	**

Ratio of income to poverty (modeled on higher ratio)	Odds ratio	95% confidence interval	*P* value statistical significance
American Indian/Alaska Native	0.968	(0.912, 1.027)	
Asian	1.870	(1.691, 2.067)	**
Native Hawaiian or Other Pacific Islander	1.485	(1.163, 1.896)	**
More than one race	1.368	(1.140, 1.642)	**
Ethnicity (reference: not Hispanic or Latino)			
Hispanic or Latino	0.677	(0.628, 0.729)	**
Marital status (reference: not married)			
Married	2.069	(1.973, 2.170)	**
Education attainment (reference: less than high school)			
High school diploma or GED	1.537	(1.375, 1.718)	**
Some college or technical training	2.640	(2.343, 2.975)	**
Associate's degree	3.468	(3.180, 3.782)	**
Bachelor's degree or higher	7.155	(6.294, 8.134)	**
Employment status (reference: not employed)			
Employed	2.694	(2.610, 2.781)	**
High school enrollment (reference: currently not enrolled)			
Currently enrolled	2.483	(2.230, 2.765)	**
College enrollment (reference: currently not enrolled)			
Currently enrolled	1.192	(1.125, 1.263)	**

Note: **Statistical significance at $p < .05$.

Summary of Results

Overall, across a 1-year interval, there were statistically significant positive effects for PAT caregivers for employment, high school enrollment, and college enrollment relative to the comparison group. This suggests that families who participate in PAT achieve better employment and education enrollment outcomes regardless of the caregiver's race or ethnicity.

Moderation

To answer the following research questions, we applied a moderation analysis to assess a differential effect of PAT related to different racial and ethnic subgroups.

- Does race or ethnicity moderate the employment and education increase over time among PAT caregivers and a comparison group?
- Are some racial and ethnic subgroups affected more than others?

Race and ethnicity were used as a moderator to assess if there is a differential effect of treatment on the employment and education enrollment for different race and ethnicity subgroups. Race and ethnicity measurement was categorized into six subgroups as follows:

- Hispanic (33.3%)
- Non-Hispanic, AI/AN (3.1%)
- Non-Hispanic, Asian and Native Hawaiian or Other Pacific Islander (AAPCHO) (2.4%)
- Non-Hispanic, Black or African American (19.5%)
- Non-Hispanic, Multiracial (1.4%)
- Non-Hispanic, White (40.4%)

Analytic Methods

We used the DID method to estimate differences within racial and ethnic subgroups for each subgroup of interest. We then took the DID estimates for each race and ethnicity subgroup and compared them to each other using a difference-in-difference-in-difference (DDD) design. The DDD allowed us to compare the outcomes of interest among racial and ethnic subgroups between two samples. We added covariates (i.e., age, gender, marital status, educational attainment, and poverty status) to control for differences in observed characteristics.

The DDD method is an extension of DID. The DDD estimator can be viewed as the difference between two DID estimators. It also can be referred to as heterogeneity of treatment effects, effect modification, or an interaction.

The DDD method compares the changes in outcomes over time in one subgroup between a treatment and comparison group, compared to the similar difference for another subgroup. The DDD estimator for the effect of treatment is

$$DDD = \left[\left(\bar{y}_{T,B,Post} - \bar{y}_{T,B,Pre}\right) - \left(\bar{y}_{C,B,Post} - \bar{y}_{C,B,Pre}\right)\right] - \left[\left(\bar{y}_{T,A,Post} - \bar{y}_{T,A,Pre}\right) - \left(\bar{y}_{C,A,Post} - \bar{y}_{C,A,Pre}\right)\right]$$

where *T* refers to the treated group, *C* refers to the comparison group
Pre refers to before treatment, *post* refers to after treatment
B refers to subgroup B, and *A* refers to subgroup A

- DDD in the regression context is:

$$Y = \beta_0 + \beta_1 Time + \beta_2 Treatment + \beta_3\, Moderator + \beta_4 Time \times Treatment + \beta_5 Moderator \times Time + \beta_6 Moderator \times Treatment + \boldsymbol{\beta_7\, Moderator} \times \boldsymbol{Time} \times \boldsymbol{Treatment} + \beta_8 Covaritate + \varepsilon$$

Model Building

To build the model, we fit a binary logistic model with a DID design to compare the changes in outcomes over time for each subgroup between treatment and comparison groups. Additionally, we fit binary logistic regression with a DDD design to compare the changes in outcomes over time in one subgroup between a treatment and comparison group, compared to the similar difference for another subgroup.

Results

Overall, race and ethnicity had a significant moderating effect on employment and education outcomes. The detailed results for employment and education outcomes are displayed in exhibits 20 and 21. Significant findings are graphically represented in exhibit 22.

Exhibit 20. Pooled estimates for employment and education outcomes, DID design

Race and ethnicity subgroup comparison	Employment status			High school enrollment			College enrollment		
	Odds ratio	95% confidence interval	*P* value	Odds ratio	95% confidence interval	*P* value	Odds ratio	95% confidence interval	*P* value
DID: Within each racial and ethnic subgroup[1]									
Hispanic	0.965	(0.854, 1.090)		1.791	(1.307, 2.455)	**	1.768	(1.326, 2.359)	**
AAPCHO	1.105	(0.702, 1.739)		2.990	(0.396, 22.55)		0.048	(0.013, 0.181)	
AI/AN	1.789	(1.196, 2.675)	**	2.038	(0.784, 5.299)		2.798	(1.360, 5.757)	**
Black	1.316	(1.124, 1.541)	**	1.904	(1.357, 2.672)	**	1.160	(0.887, 1.517)	
Multiracial	0.731	(0.397, 1.346)		1.392	(0.397, 4.874)		2.149	(0.819, 5.639)	
White	1.279	(1.143, 1.431)	**	1.354	(0.991, 1.850)	*	0.807	(0.645, 1.011)	

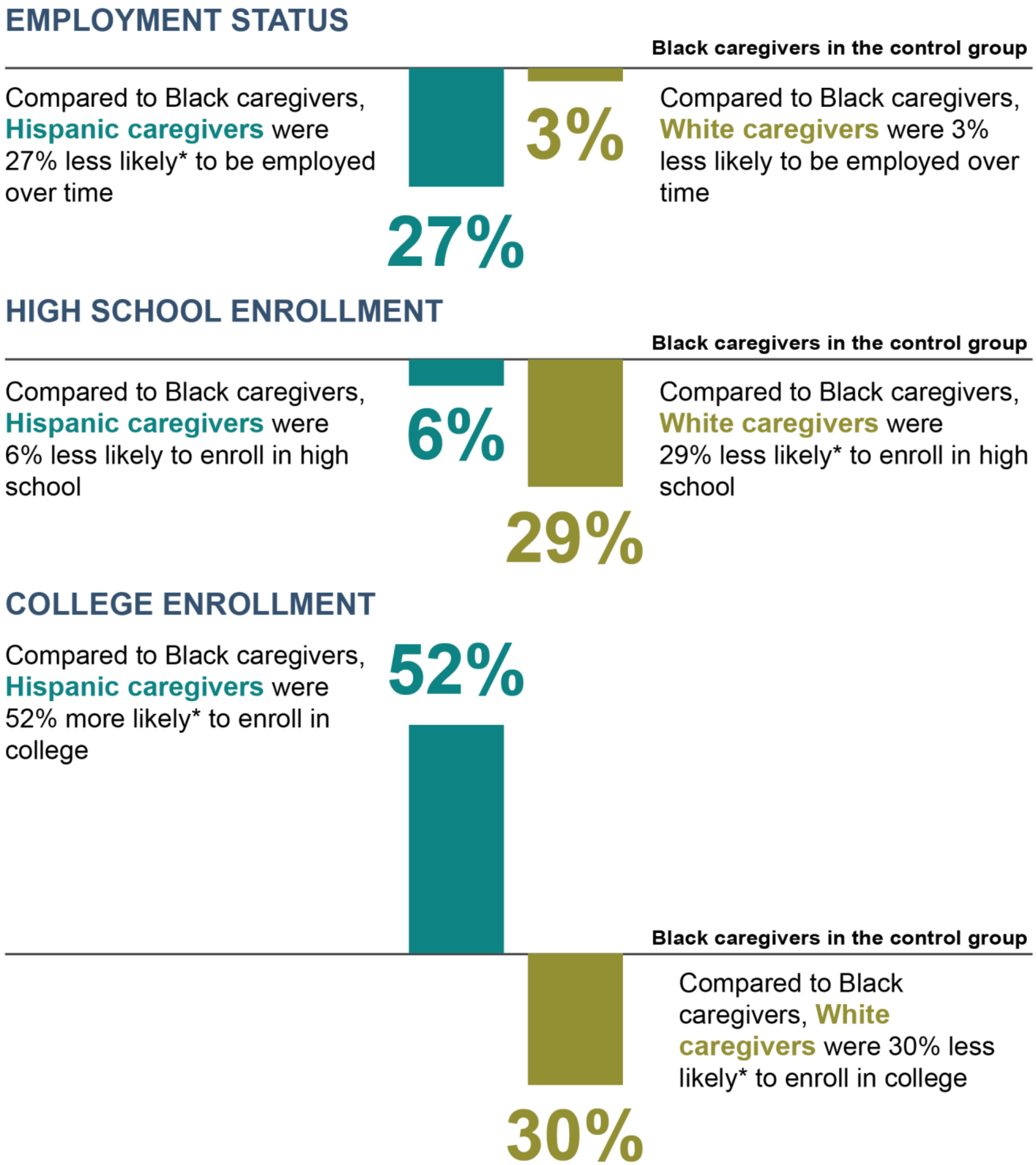
Note: *Statistical significance at $p < .1$. **Statistical significance at $p < .05$. [1]Bonferroni corrections were performed to correct for occurrence of false positives.
AAPCHO: Asian and Native Hawaiian or Other Pacific Islander; AIAN: American Indian and Alaska Native; DID: Difference in Difference

Exhibit 21. Pooled estimates for employment and education outcomes, DDD design

Race and ethnicity subgroup comparison	Employment status			High school enrollment			College enrollment		
	Odds ratio	95% confidence interval	*P* value	Odds ratio	95% confidence interval	*P* value	Odds ratio	95% confidence interval	*P* value
DDD: Between racial and ethnic subgroups									
Hispanic vs. AAPCHO	0.873	(0.616, 1.238)		0.599	(0.131, 2.736)		36.94	(12.18, 112.0)	**
Hispanic vs. AI/AN	0.539	(0.393, 0.740)	**	0.879	(0.416, 1.857)		0.632	(0.351, 1.139)	
Hispanic vs. Black	0.733	(0.632, 0.850)	**	0.941	(0.669, 1.324)		1.524	(1.120, 2.075)	**
Hispanic vs. Multiracial	1.319	(0.830, 2.097)		1.287	(0.493, 3.361)		0.823	(0.376, 1.797)	
Hispanic vs. White	0.754	(0.666, 0.854)	**	1.323	(0.954, 1.835)		2.190	(1.636, 2.933)	**
AAPCHO vs. Black	0.839	(0.586, 1.201)		1.570	(0.343, 7.190)		0.041	(0.014, 0.122)	**
AAPCHO vs. Multiracial	1.511	(0.857, 2.663)		2.148	(0.367, 12.56)		0.022	(0.006, 0.083)	**
AAPCHO vs. White	0.864	(0.609, 1.224)		2.208	(0.484, 10.08)		0.059	(0.020, 0.175)	**
AI/AN vs. AAPCHO	1.619	(1.026, 2.556)	**	0.682	(0.130, 3.585)		58.46	(17.15, 199.3)	**
AI/AN vs. Black	1.359	(0.985, 1.875)	*	1.070	(0.504, 2.274)		2.412	(1.343, 4.333)	**
AI/AN vs. Multiracial	2.447	(1.416, 4.227)	**	1.464	(0.454, 4.723)		1.302	(0.519, 3.268)	
AI/AN vs. White	1.399	(1.023, 1.912)	**	1.505	(0.713, 3.175)		3.467	(1.962, 6.125)	**
Multiracial vs. Black	0.556	(0.348, 0.888)	**	0.731	(0.279, 1.917)		1.853	(0.849, 4.045)	
Multiracial vs. White	0.572	(0.360, 0.907)	**	1.028	(0.394, 2.683)		2.663	(1.226, 5.782)	**
White vs. Black	0.972	(0.842, 1.122)		0.711	(0.506, 1.000)	**	0.696	(0.521, 0.930)	**

Note: *Statistical significance at $p < .1$. **Statistical significance at $p < .05$. AAPCHO: Asian and Native Hawaiian or Other Pacific Islander; AIAN: American Indian and Alaska Native; DDD: Difference in Difference in Difference

Exhibit 22. Select key findings from the moderation analysis



* indicates statistically significant difference between racial and ethnic subgroups (*P* < .05)

Summary

This quasi-experimental design used secondary data to explore family economic well-being outcomes for caregivers participating in PAT. The results suggest that caregivers who participate in PAT have significant increases in employment status and enrollment in high school or college relative to a comparison group and that these differences are moderated by racial and ethnic subgroups. The significant positive findings suggest that components of the PAT model may contribute to caregiver and family economic well-being.

We did not find evidence of increased family income for PAT caregivers relative to the comparison group. Many factors could affect family income, including external ones such as employment opportunities within the community. One example related to caregivers is the challenge of isolating mixed effects for caregivers' income if they enroll in school while receiving PAT services. For example, when caregivers are enrolled in school, their income may be lower if they are not working but may be higher if they are receiving financial aid. Additionally, 1 year is probably not enough time to observe an increase in education translated into increased income. Finally, only cash sources were included in the family income measurement. Cash income is highly associated with employment-based income like salary or wages. PAT caregivers may be more likely to have a short-term increase in noncash sources (e.g., Women, Infants and Children, Supplemental Nutrition Assistance Program) from referrals rather than cash sources that may be more long term.

The next step should include engaging families and parent educators in focus groups to reflect on analysis results. The groups could explore the historical context around barriers to employment and education for different racial and ethnic groups. The results could inform a theory of change about how PAT may contribute to reducing barriers within systems and institutions, thus leading to improved employment and education outcomes for families. Ultimately, precision home visiting aims to explore whether specific interventions work best for particular families and contexts and why and how they work.[5] For PAT, that could include asking questions about specific PAT interventions that may play a role in helping families overcome structural and institutional barriers to achieving their employment and education goals.

[5] https://www.hvresearch.org/precision-home-visiting/introduction-to-precision-home-visiting/

References

Douglas, F., Xiang, Z., & Zbigniew, K. (2020). Real-world health care data analysis: Causal methods and implementation using SAS. SAS Institute, 44.

Hainmueller, J. (2012). Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies. Political Analysis, 20(1), 25–46. https://doi.org/10.1093/pan/mpr025

Liu, Y., & De, A. (2015). Multiple imputation by fully conditional specification for dealing with missing data in a large epidemiologic study. International Journal of Statistics in Medical Research, 4(3), 287–295. https://doi.org/10.6000/1929-6029.2015.04.03.7

Appendix: Data Reflections

Data Quality

The following examples identify data quality issues that were addressed in the data preparation phase. These quality measures could be addressed in the future through Penelope programming (e.g., data validations upon entry) or affiliate-level data quality reports (e.g., duplicate records).

- Caregiver and/or child birth dates were outside of the valid range (e.g., child birth date used for caregiver birth date or caregiver birth date set to default).
- Enrollment and exit dates were missing or illogical (e.g., exit before enrollment date or family had not received services in a long time but had no recent hold or exit date).
- Family hold dates were illogical (e.g., family hold start date was after hold end date, family hold end date was after exit date).
- Some families had more than one Family Information Record or multiple service IDs. Different enrollments or duplicate enrollments (two enrollment dates close together) could not be determined.
- There were three versions of hold dates, and many of them were invalid; hold dates were rarely recorded.
- Stressors contained illogical (e.g., end date without a start date) and invalid (e.g., stressor end date after exit date) start and end dates.
- Average monthly and annual incomes were sometimes entered incorrectly (e.g., annual income in the monthly income field).
- Some families had more than one personal visit entered for the same day.
- Some caregivers or case records had duplicate forms that contained different responses to the same question.
- Caregiver data of interest for the study were missing.
- Information on parent educator race and ethnicity and years of experience was often missing.
- According to the PAT Data-in-Motion Manual, the Parent/Guardian Information Record and Family Information Record should be completed within 90 days of enrollment and updated at least annually thereafter. Some affiliates have affiliate procedures or funder conditions that require more frequent review and updating of items in this record. Sometimes the same entry is added with a different date to indicate the information was updated, and other times, no date is associated with an update. Because data can be collected at various time points for these forms, it is challenging to identify baseline and follow-up measures for analyses. Data cleaning revealed that these forms were often not updated according to the guidelines. The following are some examples:
  - No baseline measurement was taken within 90 days of enrollment.

    - More than one update occurred during the year (sometimes required by funders).
    - Data were updated 6 months after enrollment and no updates thereafter; we had to assume there was no change after 6 months and used the data entered at 6 months.
    - No follow-up or update of the data occurred after the baseline measure; it was unclear whether there were no updates or the follow-up was missed.
- According to the PAT Data-in-Motion Manual, there is no required timeframe for the family to set an initial goal, but at least one goal is typically set within 90 days of enrollment. The PAT Goal Record is started when a goal is first set and is updated as progress is made toward the goal. PAT recommends that parent educators document a progress note on the Goal Record at least monthly until the family has met the goal or has decided to no longer work on it. Because data are not required at intake and can be collected at various time points for this form, it is challenging to identify the goals set up at intake.
- In the Goal Record form, the parent educator can indicate to whom the goal pertains. The goal can be documented under an individual but pertain to another family member or to several family members. This makes it challenging to tie the goal record to a specific individual.
- There is no way to tie the achievement of a goal to updated information on other forms in Penelope (e.g., update the education status when a family is enrolled in school).

Data Collection Opportunities

Based on observed data quality issues, here are some considerations for future Penelope system and variable modifications that would enhance data availability for future studies.

Affiliate Sites

- Consider adding affiliate details to Penelope, including the state program, funding source(s), dates of funding sources, curriculum subscribers only, and research sites (e.g., Family and Child Education study).
- For geographic analyses, zip code data are more useful than self-identified family community data.

Caregivers

- Add a unique ID for all caregivers and case files (i.e., one ID that links all records across data files and is present in all data files).
- Add or activate a date or time stamp for all documents completed in Penelope.
- Add an enrollment ID to track the reenrollment for caregivers.
- Designate a primary caregiver on the case file or individual record.
- Expand individual gender response options (e.g., female, male, transgender, nonbinary) to allow for disaggregating the data.

- Prenatal enrollment designation should be collected separately (i.e., prenatal is a response option for “individual gender: male/female/prenatal”).
- Some stressors have more than one version (e.g., low income options a, b, c and insecure housing a, b, c) and could be merged.
- Consider what constructs are important for examining the types of families that PAT serves, which would allow for latent class analysis. This could allow the delivery of services to be tailored depending on families’ different needs.

Expected Visit Frequency

- Two stressors may not be an appropriate cutoff for identifying high-need families depending on the type of stressors identified (e.g., some are more severe than others).
- In the current calculation, all stressors carry the same weight. Consider weighting more severe stressors to create tiers.
- Some stressors may not need end dates because they should not change over time. The start date could be recorded at baseline or when stressor data are collected (e.g., parent is or was incarcerated, death in the immediate family, parent or guardian is a survivor of intimate partner violence.
- Number of children in the family could affect the number of personal visits (i.e., families with multiple children may need more personal visits).

Goals

- Consider tracking goals by individual rather than by the case record.
- Consider adding more specific goal options in the existing quantitative question about goal areas. The current 11 broad goal areas (e.g., education) do not give enough information about the goal details (e.g., enroll in high school, complete a degree) in a quantitative format for descriptive analyses.

Parent Educators

- Consider converting qualitative parent educator fields to quantitative fields (e.g., field of study is an open-ended response but could be coded).
- Consider collecting additional data on parent educators (e.g., years of experience, training, caseload).
- Consider the ideal language match between families and parent educators. For example, is it important that the parent educator have some level of fluency in Spanish, or should they be fluent?