

Linking Academic Achievement To
Attendance in Georgia's 21st Century
Community Learning Centers Program
A Propensity Score Demonstration

GSU Evaluation Team Members

- Jeremy Lingle, Ph.D. Candidate

Jlingle1@gsu.edu

- Sheryl Gowen, Ph.D.
- Carolyn Furlow, Ph.D.
- Syreeta Skelton, M.A.
- Corrie Davis, Ph.D.



Background and Purpose

The Georgia Department of Education contracted with Georgia State University to evaluate the 21st Century Community Learning Centers on student academic achievement.

This paper presents procedures for using propensity score matching to test the hypothesis that the relationship between test scores and program attendance is mediated by the academic focus of the after school program.

21st CCLC Program

- Purpose of 21st CCLC program
 - Increase the percentage of youth who meet expectations on statewide assessment test
 - To target youth who attend Title 1 schools
 - To target youth who are struggling academically
- Service Delivery
 - More than 25,000 students each year
 - 48 programs operating 205 centers state-wide

Research Overview

- Research Purpose
 - Secondary analysis of the data to further explore the average treatment effect on math
 - Test hypothesis that relationship of attendance and the math outcome is mediated by the level of math focus of the programs
- Quasi-experimental design
 - Construct treatment and control groups from data set of 21st CCLC attendees
 - Identify individuals from the treatment and control groups that are comparable
 - Determine average difference on outcome measure between matched pairs
 - Identify optimal academic focus of programs for improving outcome

Evaluation Challenges

- Part 1: The Outcome Measure
- Part 2: Comparison Groups

Part 1: The Criterion Referenced Competency Test (CRCT)

- Tests are not vertically scaled
 - Unable to model outcomes across grades using scale score
- Outcome scores are reported as ordinal rather than cardinal
 - Exceeds Expectations
 - Meets Expectations
 - Does Not Meet Expectations

Part 2: Defining Comparison Groups

- Treatment group and Control group design
- Random Assignment presented both political and logistic challenges
- *Selection Bias*: Comparison of after school program attendees to general population of feeder schools
 - Likely difference in a number of factors, notably academic need
- *Solution*: Limit sample to only youth who attended the programs
 - Quasi-experimental design
 - Treatment group = Youth who attended 30 days or more
 - Control group = Youth who attended from 1 to 29 days

Matching Group Members

- *Selection Bias*: Comparison of treatment group to control group
 - Imbalances discovered in the following demographic categories
 - Ethnicity (e.g., proportion African American, Caucasian)
 - Free/Reduced Lunch Status (proxy for SES)
 - Grade level
- *Solution*: Matching on covariates
 - Challenge of Matching per covariate
 - Number of cells required
 - Ethnicity x Lunch Status x Grade level
 - $(2 \times 2) \times 2 \times 8 = 64$ cells
 - Matching on Propensity Scores

Creating a Comparison Group

- Propensity Scores: The conditional probability of receiving treatment.
 - In simple random assignment, the true propensity score is equal to 0.5
 - All student-level covariates appearing in the sample will have an equal probability of appearing in both the treatment and control groups
 - In quasi-experimental designs, the propensity of receiving treatment can be estimated from the measured variables available.

Balancing with Propensity Scores

- Individuals with the same estimated propensity score also have the same measured variables, regardless of treatment assignment
- Matching treatment group members with control group members on their propensity score creates balance between measured covariates
- Propensity scores methods commonly used are:
 - Matching
 - Stratification
 - Covariate adjustment/weighting

METHODS AND RESULTS

Initial Sample Demographics:

Georgia's 21st CCLC 2005-2006

- $n = 4,596$
 - 50% Female
 - 87% Free/reduced lunch
 - 11% Special needs
 - 7% LEP
 - Grade Level
 - 71% Elementary School
 - 29% Middle School
 - Ethnicity
 - 70% African American
 - 26% Caucasian
 - 4% Latino/a
 - CRCT Math Scores in Year One
 - 11% exceeded standards
 - 58% met standards
 - 32% did not meet standards

Estimating Propensity Scores

- Propensity scores were estimated through logistic regression
 - Outcome: Treatment assignment (Treatment = 1)
 - Predictors: Ethnicity, Free/Reduced Lunch Status, Gender, LEP, Special Needs/Disabilities, Grade Level, and CRCT Math Score in Previous Year.
 - Estimates the probability (as logit) of being in the treatment group, which is the propensity score.

Propensity Score Estimation

Predictor	Coefficient	SE	Odds Ratio
Intercept	0.82*	0.24	2.28
CRCT Math	0.02	0.07	1.02
Special Needs	-0.06	0.18	0.95
Female	0.12	0.08	1.23
Latino/a	0.24	0.27	1.27
African American	0.59*	0.12	1.81
Free/Reduced Lunch	0.17	0.15	1.18
LEP	-0.21	0.33	0.81
Grade Level	-0.22*	0.03	0.81

Interpretation of Coefficients

- African American youth were 1.81 times more likely to be in the treatment group than their Caucasian peers.
 - $0.82 + 0.59 = 1.41$
 - $\text{Exp}(1.41) / (\exp[1+1.41]) = 0.80$
 - An African American youth who is male, does not meet expectations on the math outcome, who is in the first grade has an 80% probability of being in the treatment group.
 - A youth who only differs in that he is Caucasian has a 69% probability of being in the treatment group.
- A one unit increase in grade level corresponds to a 19% decrease in the odds of being in the treatment group
 - $100\% * (\text{Odds ratio} - 1)$

Matching on Propensity Score

- 21st CCLC sample
 - Treatment group = 59%
 - Control group = 41%
- Nearest neighbor matching with replacement of controls (1:k)
- Matched within programs
 - To balance unmeasured group-level covariates
 - To maintain multilevel structure of data

Matched Sample

- Initial Sample = 2,693 treatment: 1,903 control
- Matched Sample = 1,366 pairs
 - 51% of the treatment youth had acceptable matches in the control group within their programs
- Contributors to the loss of sample:
 - The greater the difference between treatment and control groups on essential variables, the greater the separation on propensity scores and fewer the matches
 - Limiting to within program matches greatly diminishes the pool of controls, but necessary to maintain multilevel structure

Bias Reduction

Covariate	$d_{initial}$ ($n=4,596$)	$d_{matched}$ ($n=2,732$)	Bias Reduction
African American	13%	2%	88%
Caucasian	9%	0.4%	96%
Free/Reduced Lunch Status	14%	1%	94%
Grade Level	49%	0.4%	99%

- d is the standardized difference between the proportion/percentage of each covariate between the treatment and control groups

Research Question 1: What was the Average Treatment Effect across Programs?

- Pair-difference in outcome measure: Math CRCT 2006
 - 1 = Does not meet expectations
 - 2 = Meets expectations
 - 3 = Exceeds expectations
- Average Pair Difference = 0.119 ($sd = 0.69$)
- Effect Size, Cohen's $d = 0.17$

Research Question 2: Does math focus mediate relationship of attendance and math outcome?

- Calculated the math focus of each program:
 - (*a*) Number of activities attended by each youth that had a math-focus
 - (*b*) Number of activities attended by each youth overall.
 - Average of a / b
- Across the programs ($n = 64$), the math-focus proportion ranged from 0 to 0.88 with a mean of 0.25 ($sd = 0.18$).
- Stratified the sample into quintiles based upon the math-focus proportion.

Average Treatment Effect per Math-Focus Quintile

Quintile (range)	n_{pairs}	Mean Difference	sd	Effect Size
1 (0 - .16)	254	0.051	0.666	0.08
2 (.167 - .235)	263	0.064	0.699	0.09
3 (.242 - .29)	309	0.129	0.740	0.17
4 (.291 - .323)	279	0.168	0.707	0.24
5 (.324 - .88)	261	0.176	0.638	0.28

- Findings:
 - Benefits evident beginning at quintile 3
 - Benefits are at maximum in quintiles 4 and 5

Conclusions and Limitations

- A positive relationship exists between math outcomes and attendance in 21st CCLC programs. This relationship is mediated by the math focus of the programs.
- Youth evidenced greatest math benefits when at least 29% of the program activities had a math focus.
- Limitations
 - No control for unobserved student-level covariates
 - Loss of sample through conservative matching technique

Further Reading on Propensity Scores

- Caliendo, M. & Kopeinig, S. (2008). Some practical guidance for implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31-72.
- Hahs-Vaughn, D. L. & Onwuegbuzie, A. J. (2006). Estimating and using propensity scores with complex samples. *The Journal of Experimental Education*, 75(1), 31-65.
- Luellen, J. K., Shadish, W. R., & Clark, M. H. (2005). Propensity scores: An introduction and experimental test. *Evaluation Review*, 29(6), 530-558.
- Rosenbaum, P. R. (1986). Dropping out of high school in the United States: An observational study. *Journal of Educational Statistics*, 11(3), 207-224.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41–55.

Jeremy Lingle

Jlingle1@gsu.edu

THANK YOU!