Exploring Approaches to Constructing Comparable Groups Within Hierarchies

Doreen Finkelstein
The College Board

Presentation at the Annual Meeting of the American Educational Research Association, San Diego, CA, April 15, 2009

The Challenge: Constructing Comparable Groups for an Evaluation with Limited Data

- Program implemented school-wide (no comparison students within a school)
- Decision to implement the program was based on school-level variables (lower school achievement and serving minority and low-SES students)
- Outcome measures of interest were student-level (attitudes and beliefs about college)

School Data

- Collected student-level data from 8 schools: 5 program schools and 3 comparison schools

School	*n*	School Ach. Score	Free/ Reduced Lunch	Minority
Prgm A	35	469 (7)	34% (2)	19% (1)
Prgm B	48	430 (5)	62% (6)	99% (8)
Prgm C	22	384 (3)	52% (4)	89% (4)
Prgm D	19	500 (8)	50% (3)	66% (3)
Prgm E	22	405 (4)	59% (5)	95% (6)
Comp A	40	350 (1)	68% (7)	99% (7)
Comp B	53	442 (6)	33% (1)	63% (2)
Comp C	38	376 (2)	74% (8)	92% (5)

(#) = rank, 1-8

Variables

- School-level:
  - School Achievement Score
  - % Free/Reduced Lunch
  - % Minority
- Student-level:
  - Gender
  - Ethnicity
  - Parent Education Level*
  - Number of College Preparatory Exams Taken*

* Some hs, hs grad, some college, AA, BA, graduate degree

**PSAT/NMSQT® , PLAN ®, SAT ®, ACT ®, or Advanced Placement®

CollegeBoard
inspiring minds™

Student Data

- Student demographics, program vs. comparison:

Variable		Program Students *n* = 146	Comp. Students *n* = 131	Total *n* = 277
Gender	Male	41.1%	50.4%	45.5%
	Female	58.9%	49.6%	54.5%
Ethnicity	% Black	52.7%	57.3%	54.9%
	% Hispanic	19.9%	20.6%	20.2%
	% White	22.6%	18.3%	20.6%
Parent Educ.	(1-6)	2.81	2.86	2.83
# Exams	(0-5)	.610	.687	.646

CollegeBoard
inspiring minds™

How to Match?

- Propensity score matching (Rosenbaum & Rubin, 1983) uses logistic regression to combine and weight variables into one score
- Calculate “propensity” to be in the treatment group: propensity score variables are related to the treatment decision, control for a priori bias in outcome measures
- But, problematic to match solely on school:
  - Data probably not a random sample at each school
  - Outcome measures of interest were student-level
- Not enough data for a multi-level logistic model

Three Variations:

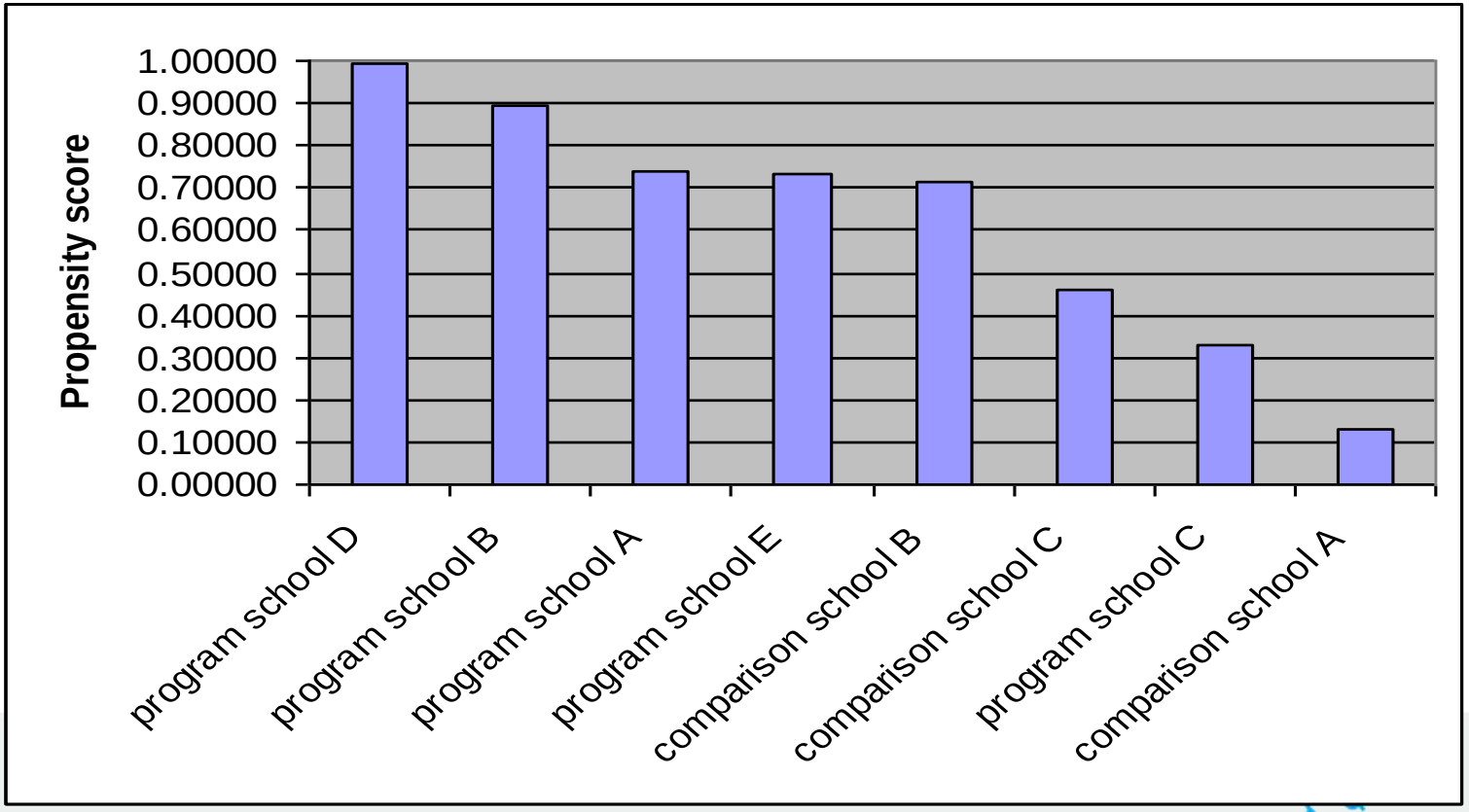
- Match on schools, but include a student-level variable
- Match on students, but include a school-level variable
- Match on schools, then match on students within those schools

Match on Schools, Include Student-Level Variable

- Use the 3 school-level variables: school achievement score, % free/reduced lunch, % minority
- Use the student-level variable most highly correlated with outcome: # of college preparatory exams taken
- With logistic regression, calculate the propensity score for each school
  - Model predicted 4/5 of program group membership, 2/3 of comparison group membership

Results of School Match

- Program Schools A and E matched to Comparison School B
- There were no other matches (within ¼ of sd of propensity score, Cochran & Rubin, 1973)



CollegeBoard
inspiring minds

Match on Students, Include School-Level Variable

- Use the 4 student-level variables: gender, ethnicity, parent education level, # of college preparatory exams taken
- Use the school-level variable most highly correlated with group membership: school achievement score
- With logistic regression, calculate propensity score for each student
  - Model predicted 80.8% of program group membership, 72.5% of comparison group membership
- 16 students could not be matched (no comparison students within ¼ sd of propensity score); after four matching runs (allowing replacement), 95% of remainder matched

CollegeBoard
inspiring minds™

Results of Student Match

- 124 program students matched, from all program schools
- Most matches came from Comparison School B

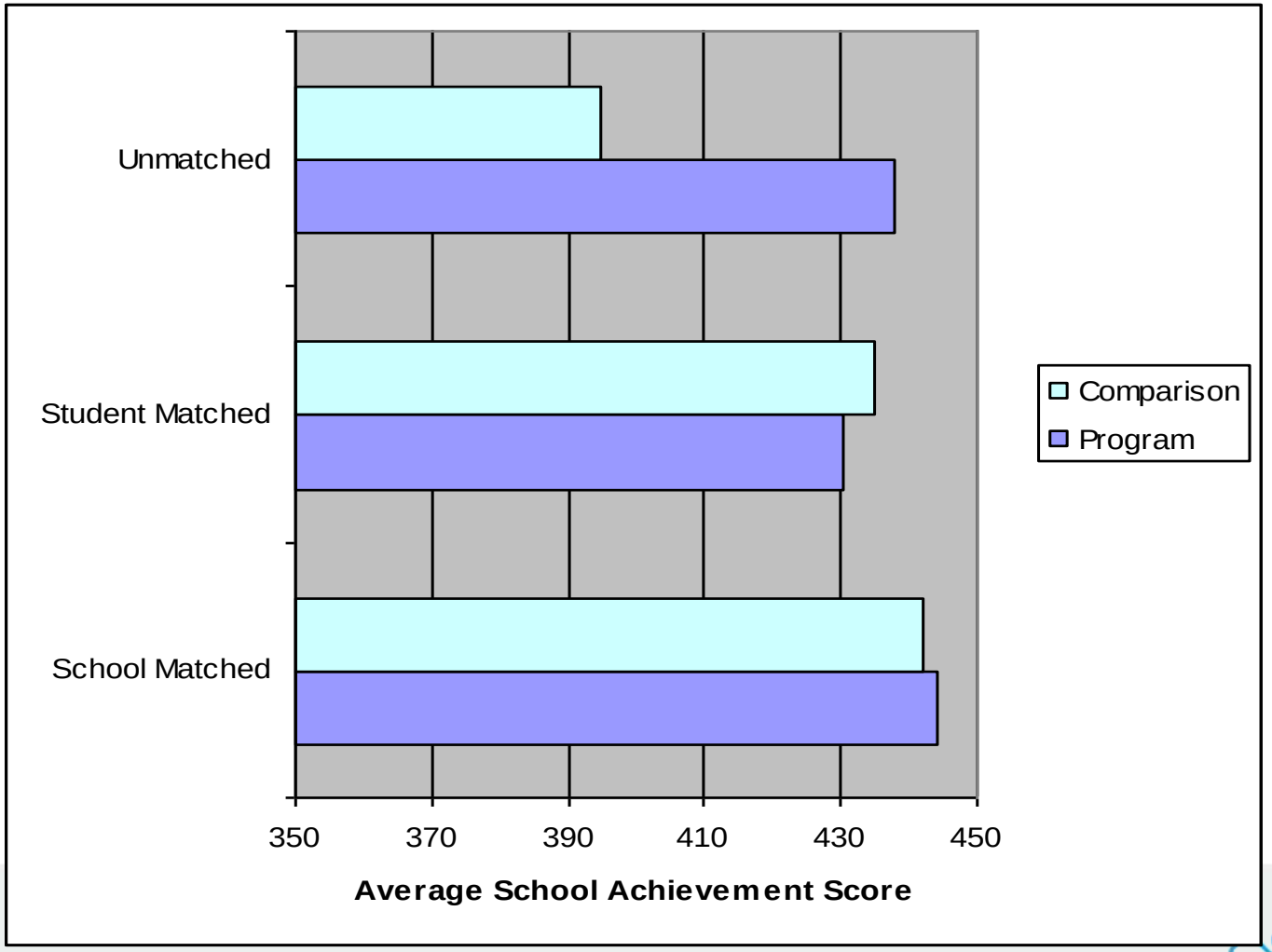
# of Matches	Comp A	Comp B	Comp C	Total
One	1	11	10	22
Two	0	9	1	10
Three	0	14	0	14
Four	0	10	0	10
Total # of Comp Students	1	44	11	56

CollegeBoard
inspiring minds

Match on Schools, Then Match on Students

- Results the same as matching on schools with student-level variable: Program Schools A and E matched to Comparison School B
- 57 program students and 53 comparison students
- Because of low match rate, no further match undertaken

Comparison of Match Results: School Achievement Score (School-Level)

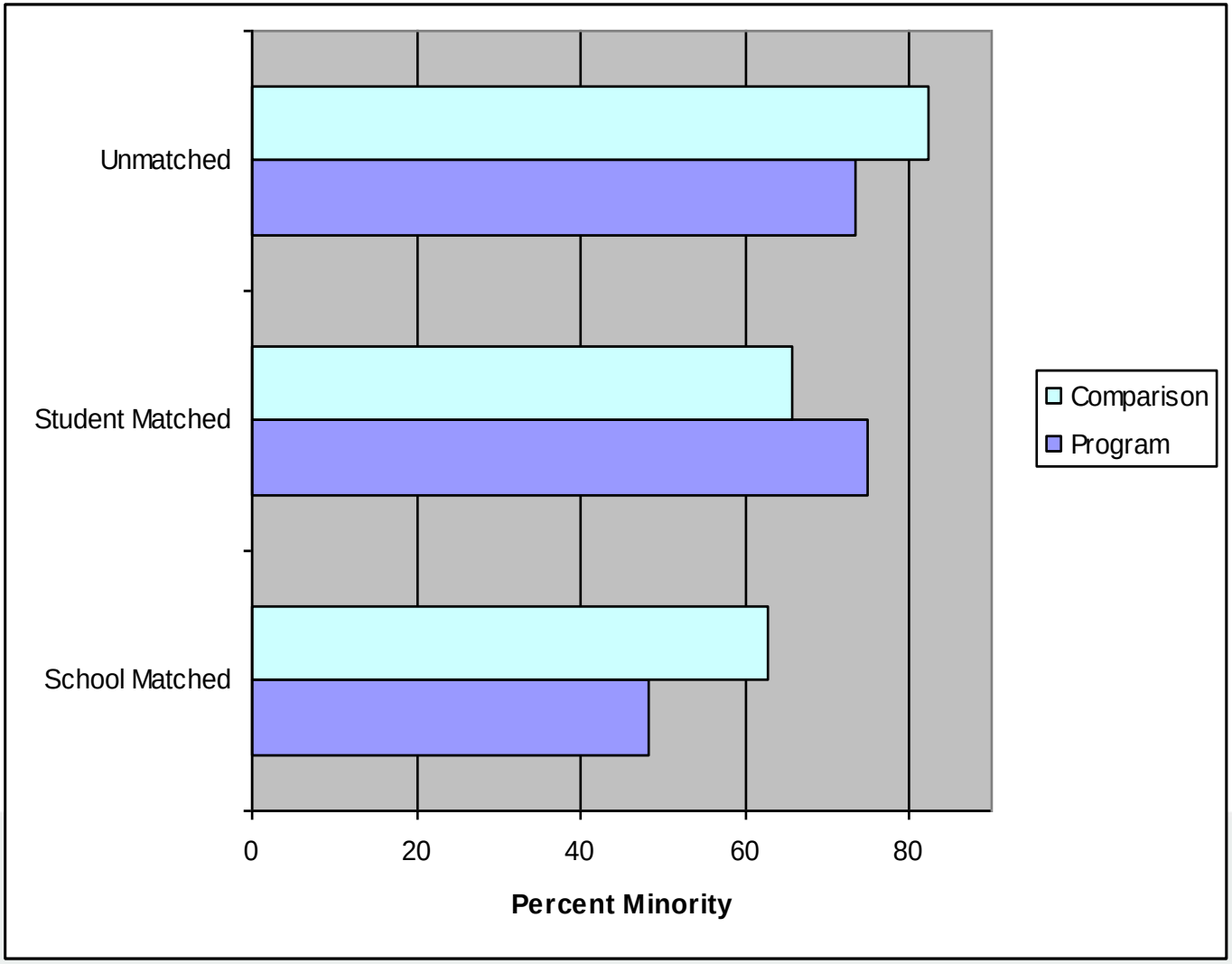


CollegeBoard
inspiring minds

Comparison of Match Results: % Free/Reduced Lunch (School-Level)

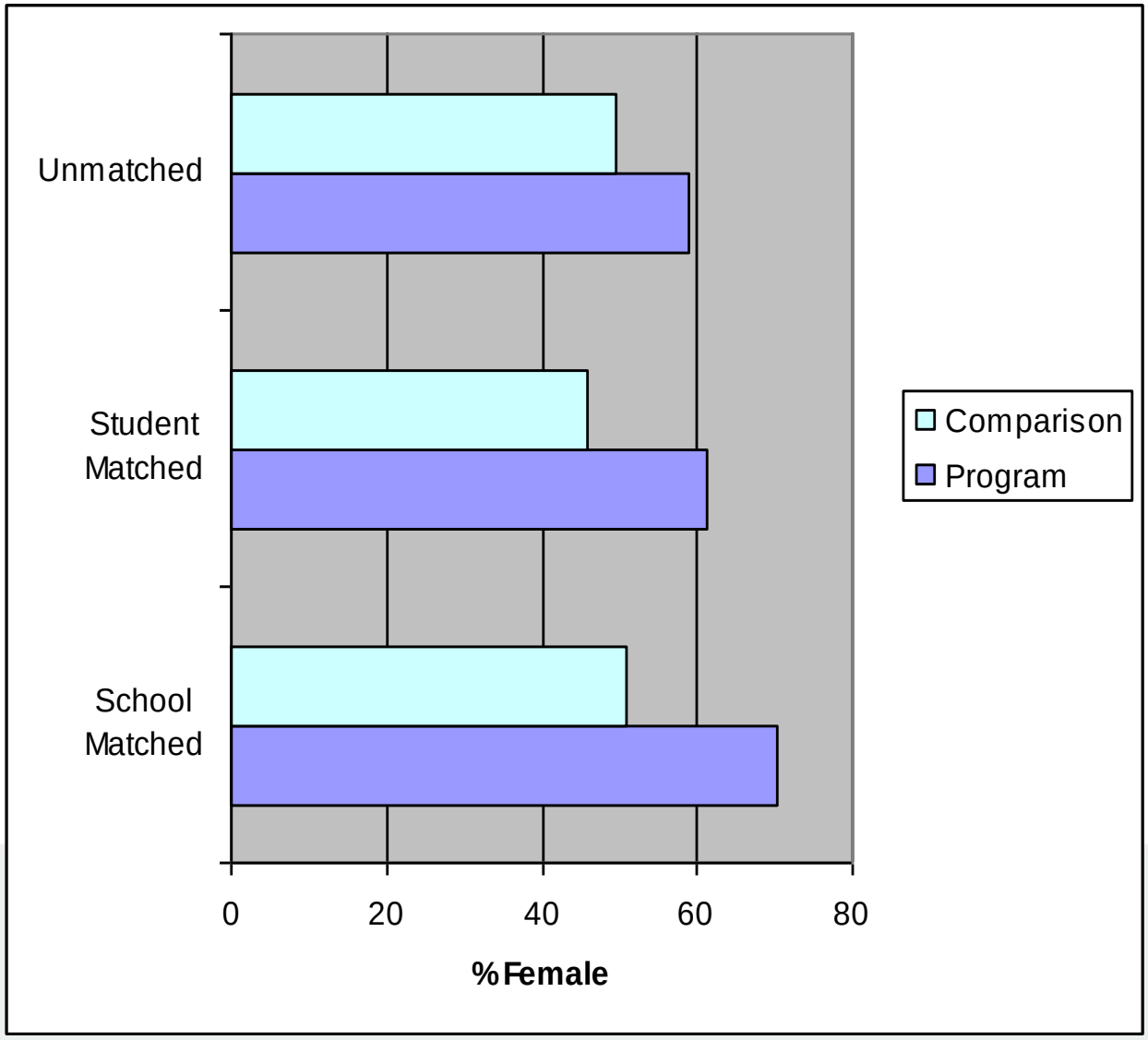
CollegeBoard
inspiring minds™

Comparison of Match Results: % Minority (School-Level)



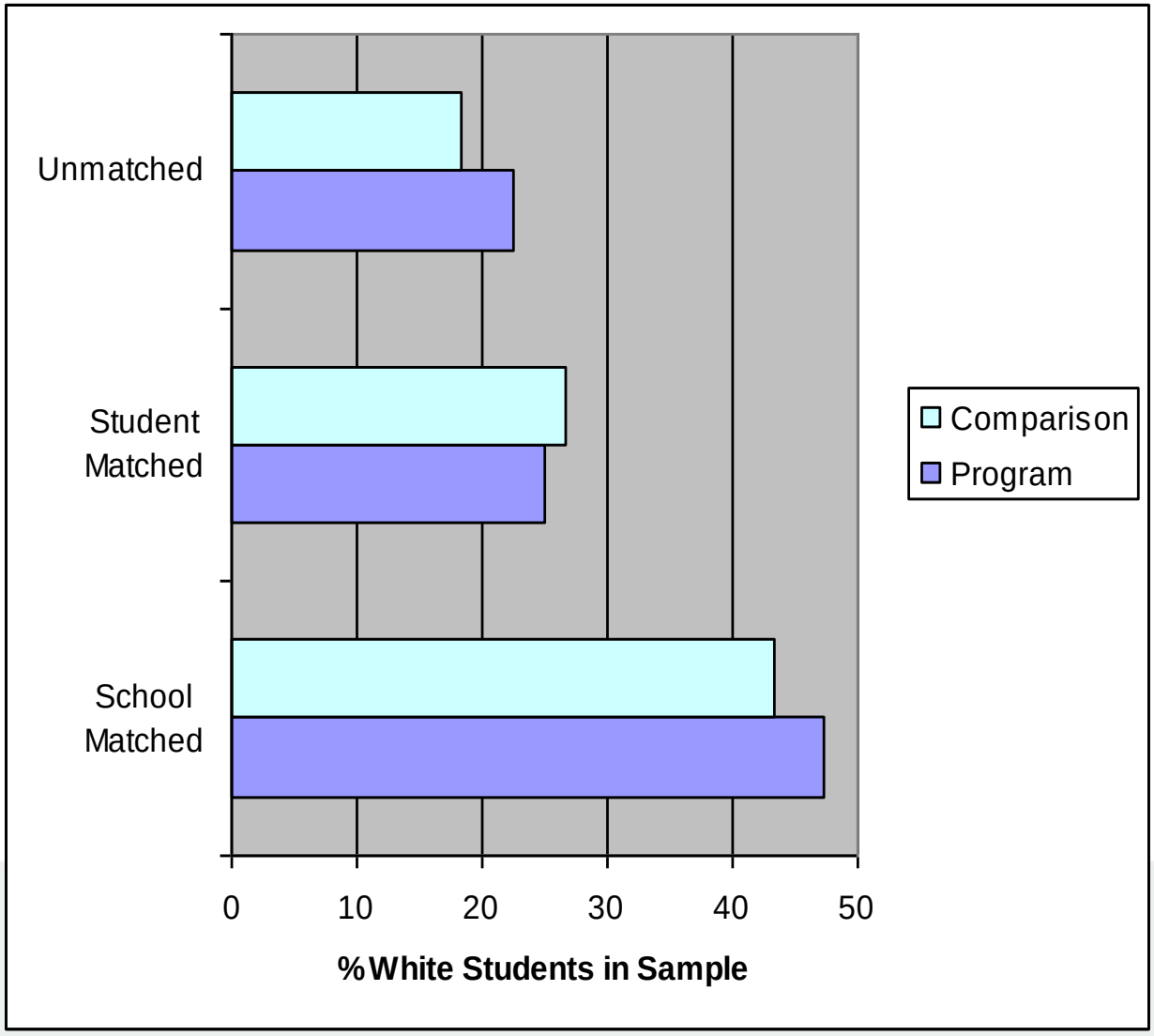
CollegeBoard
inspiring minds

Comparison of Match Results: Gender (Student-Level)



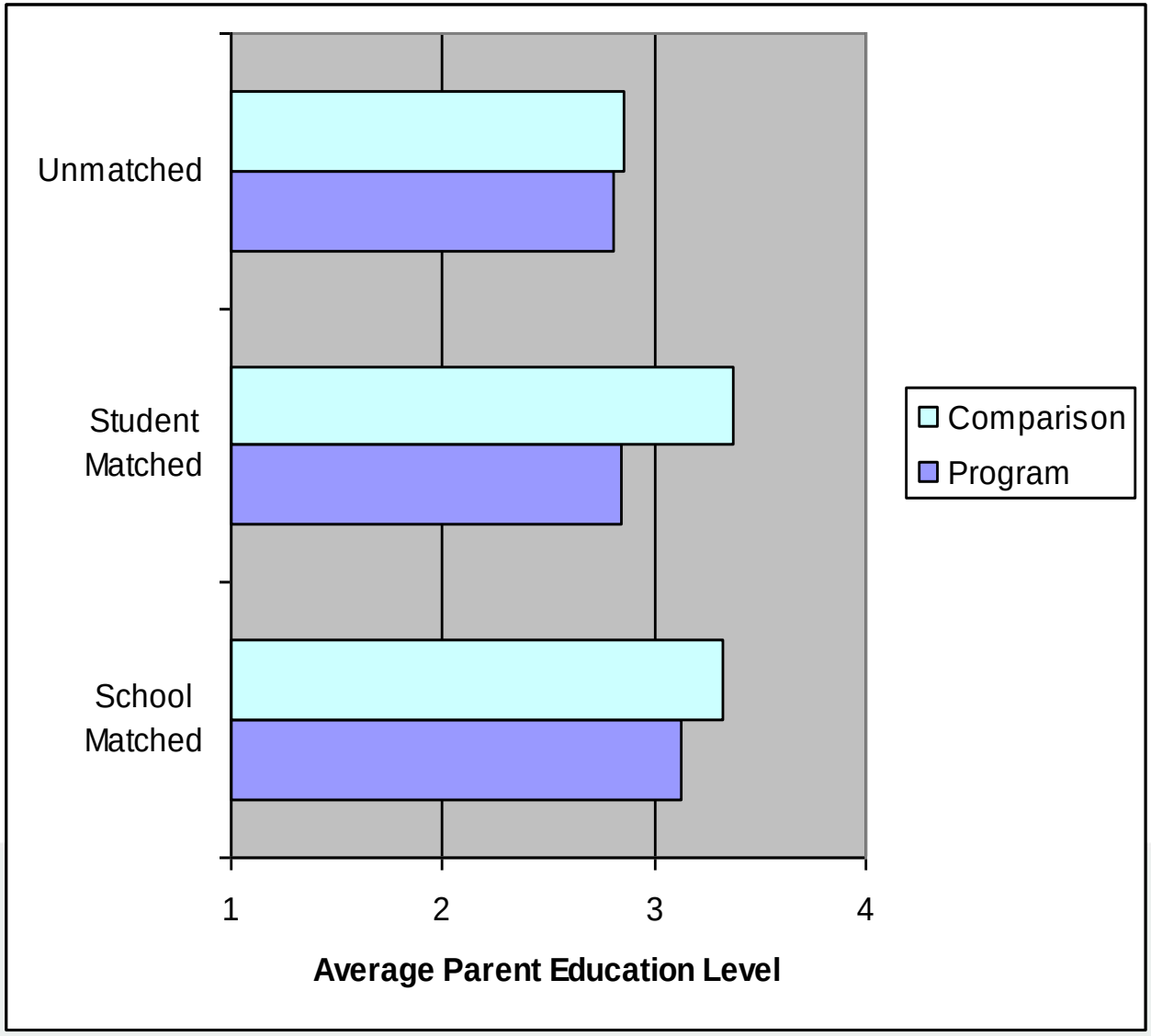
CollegeBoard
inspiring minds

Comparison of Match Results: Ethnicity (Student-Level)



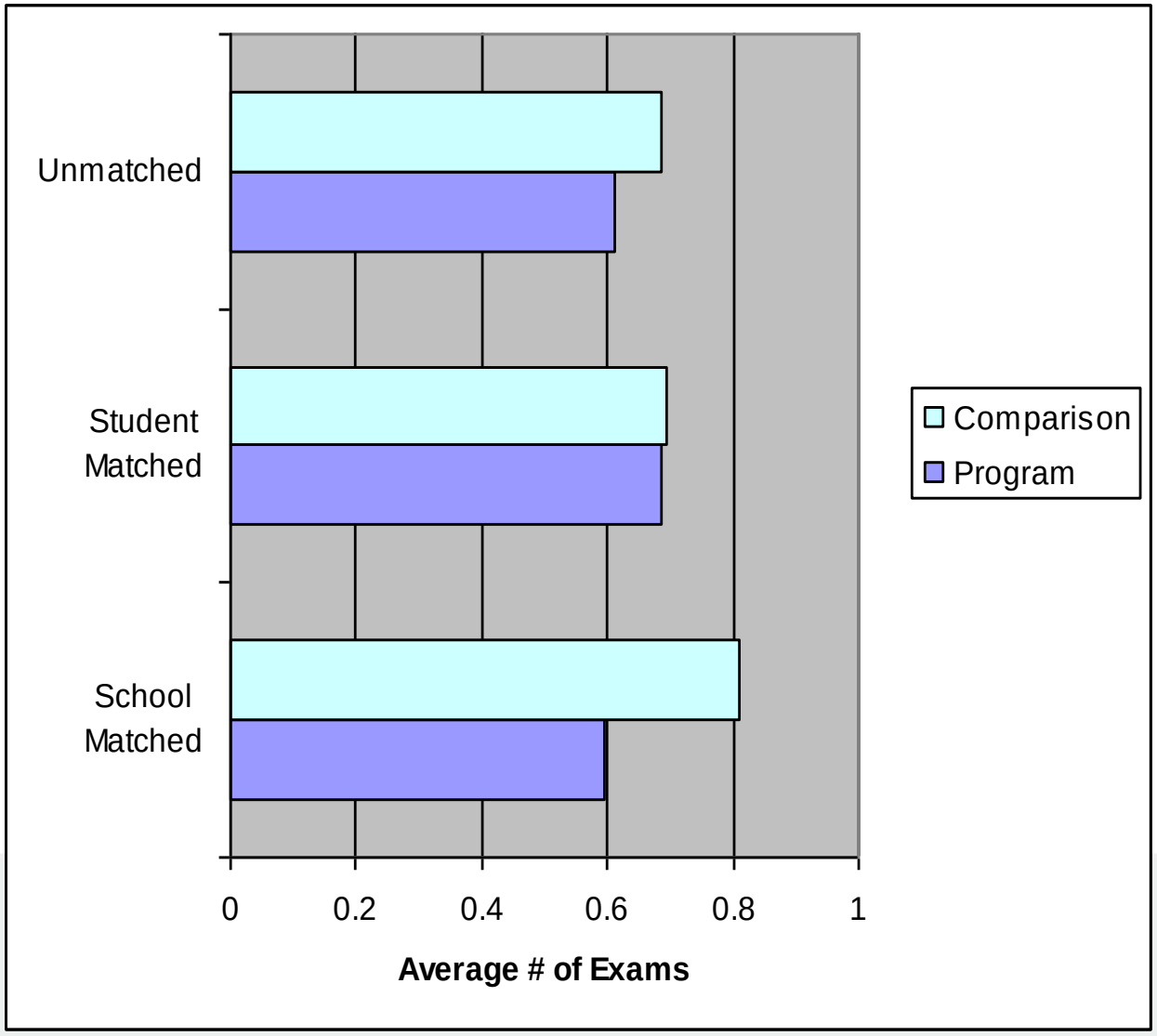
CollegeBoard
inspiring minds

Comparison of Match Results: Parent Education Level (Student-Level)



CollegeBoard inspiring minds™

Comparison of Match Results:
of College Prep Exams Taken (Student-Level)



CollegeBoard
inspiring minds™

Discussion

- Because of data limitations, no ideal approach for match
- Variations of the match attempted to take into account both school-level variables (basis of program decision) and student-level variables (to control for bias on outcome measure)
- Best match obtained by matching on students and using a school-level variable
- Possible future directions with a larger data set:
  - Use more variables
  - Compare against matching on schools, then students (as originally intended)
  - Compare against multi-level logistic model

Questions, Comments, Suggestions

- Researchers are encouraged to freely express their professional judgment. Therefore, points of view or opinions stated in College Board presentations do not necessarily represent official College Board position or policy.
- Please forward any questions, comments, and suggestions to:

  Doreen Finkelstein at: dfinkelstein@collegeboard.org

CollegeBoard
inspiring minds