MLDS Center

Maryland Longitudinal Data System

Better Data • Informed Choices • Improved Results

Dual Enrollment in Maryland: Using Propensity Scores to Strengthen Program Evaluation with State Longitudinal Data

Angela K. Henneberger
Alison Preston
Heath Witzen
Maryland Longitudinal Data System Center

2017 NCES STATS-DC Data Conference

August 3, 2017

https://mldscenter.maryland.gov/

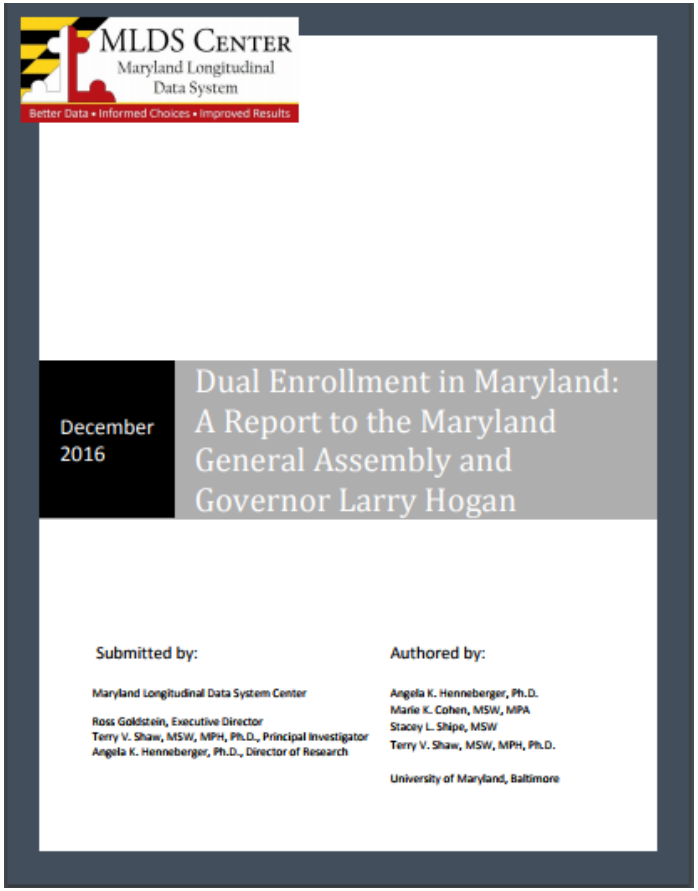
MLDS Center
Maryland Longitudinal Data System

Maryland's Dual Enrollment Report

- Annual report on dually enrolled students required by College and Career Readiness and College Completion Act (CCR-CCA) of 2013.
  - Requires the Maryland Longitudinal Data System Center (MLDSC) to report to the Governor and General Assembly:
    - the number of dually enrolled students and
    - the courses taken by dually enrolled students

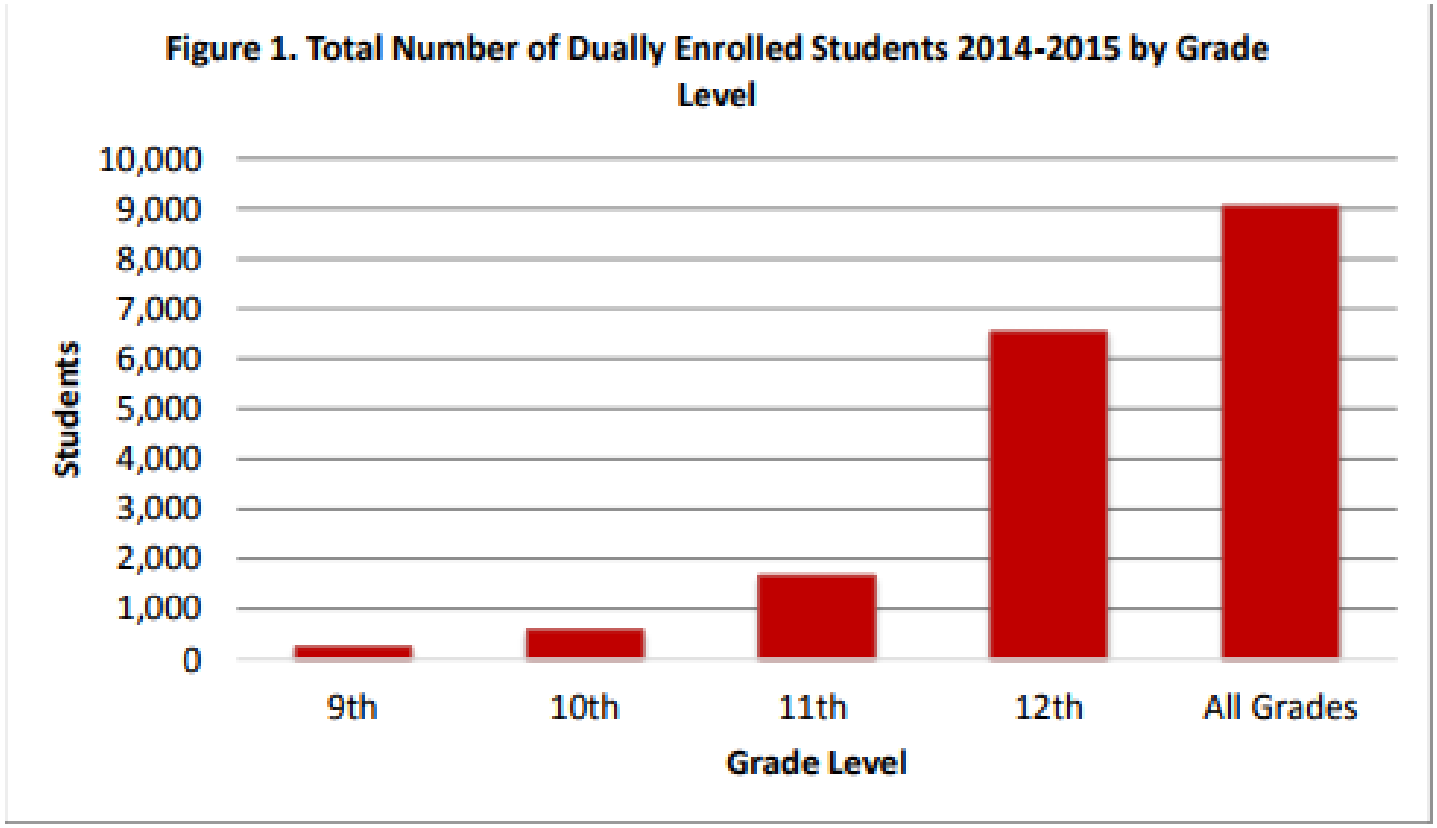
    (Education Article §24-703.1).
- A dually enrolled student is a student enrolled in both a secondary school and postsecondary institution in Maryland (Education Article §18-14A-01).

Maryland's Dual Enrollment Report



https://mldscenter.maryland.gov/DualEnrollment.html

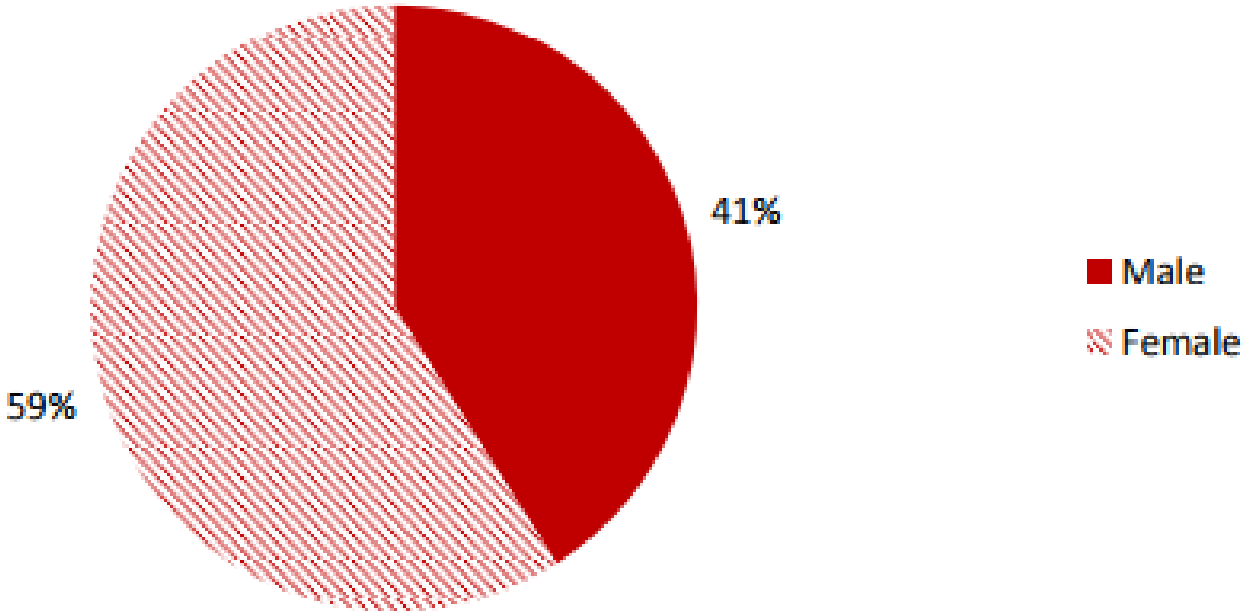
Dual Enrollment in Maryland



Source: Henneberger, Cohen, Shipe, & Shaw, 2016

Dual Enrollment by Gender

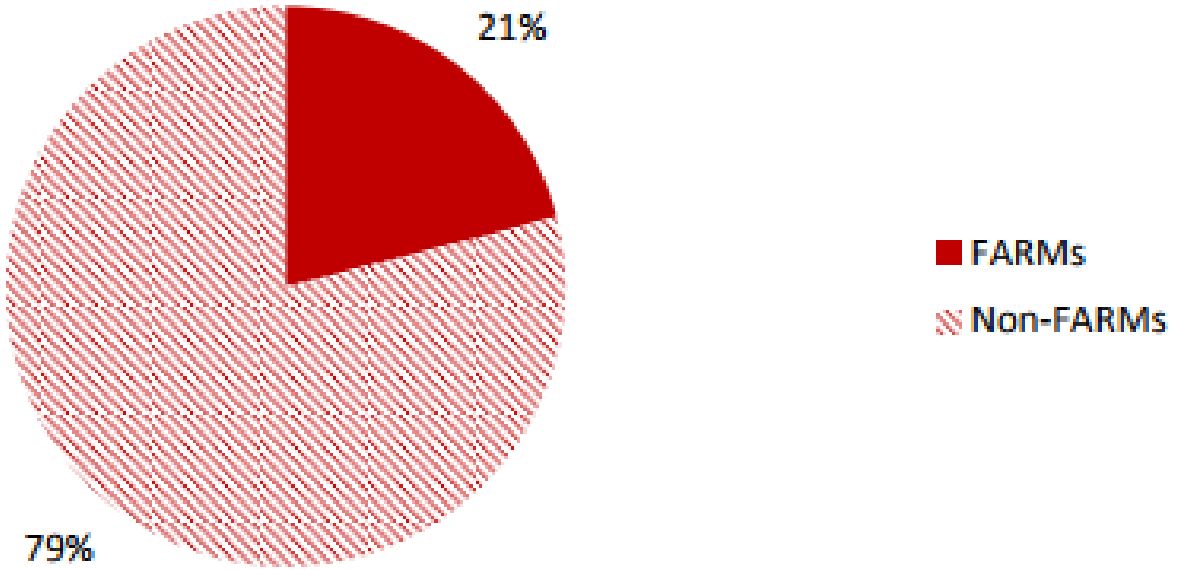
Figure 5. Gender of Dually Enrolled 12th Grade Students (2014-2015)



Source: Henneberger, Cohen, Shipe, & Shaw, 2016

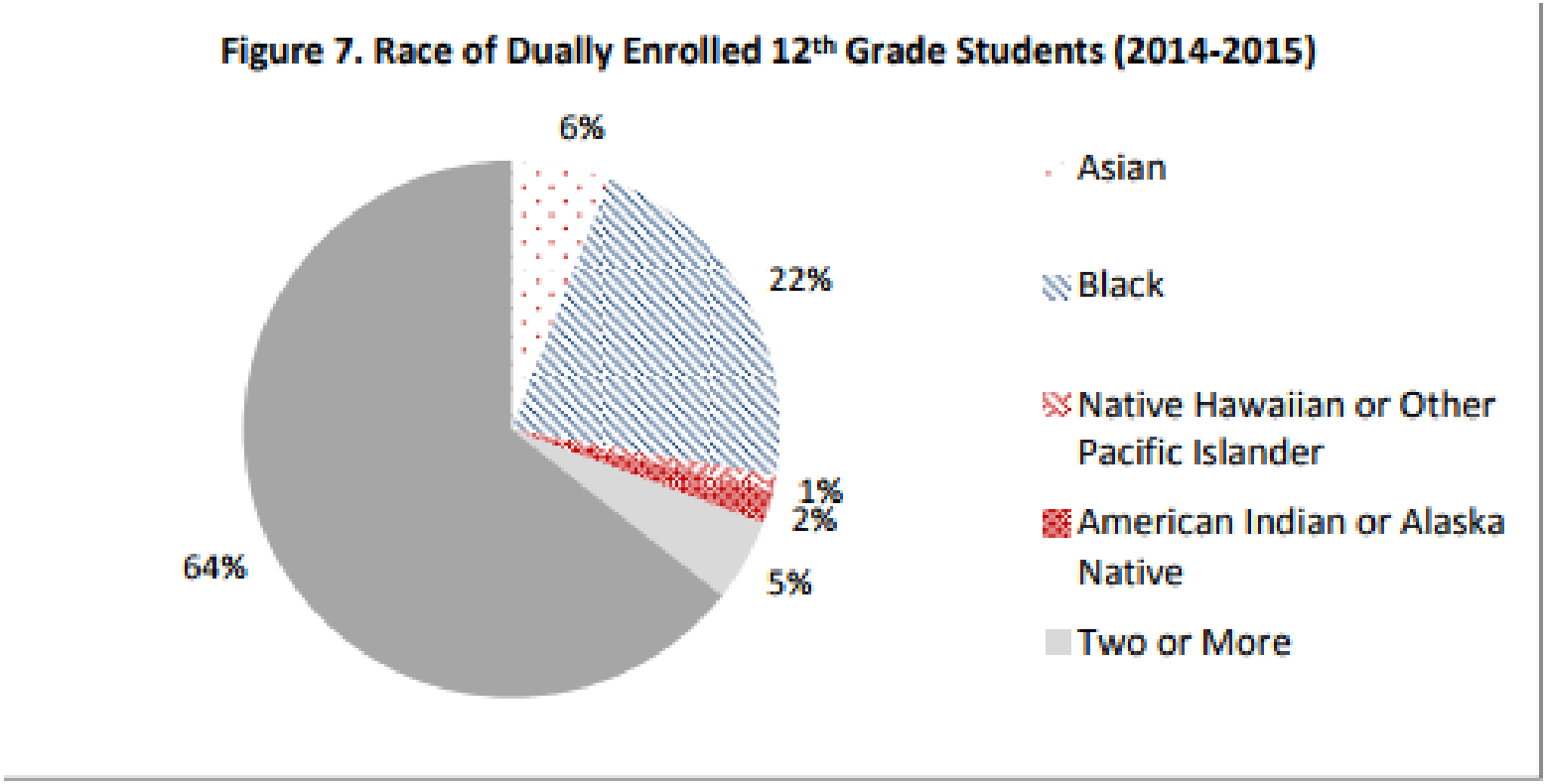
Dual Enrollment by FARMs

Figure 6. Eligibility for Free and Reduced Price Meals (FARMs) Status for Dually Enrolled 12th Grade Students (2014-2015)



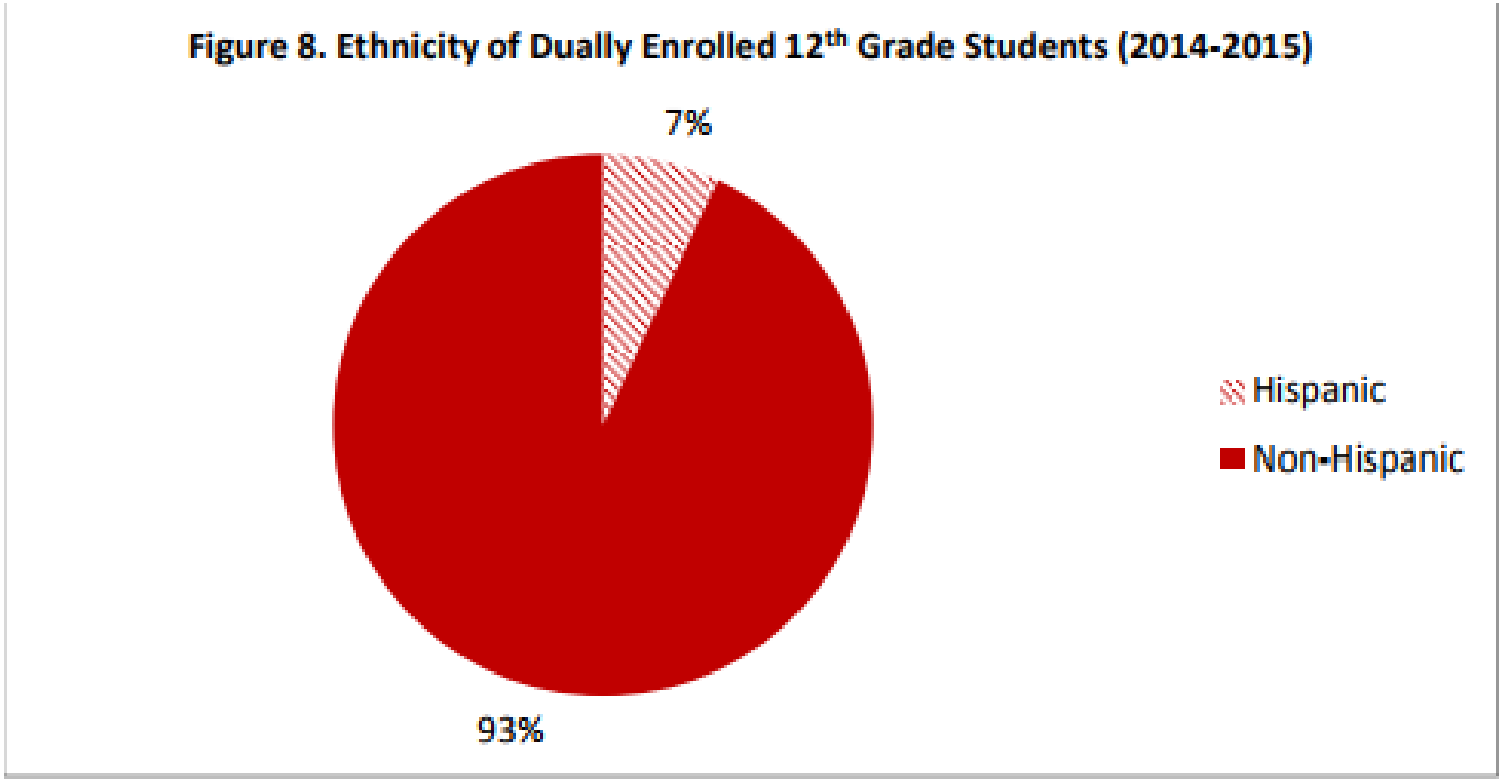
Source: Henneberger, Cohen, Shipe, & Shaw, 2016

Dual Enrollment by Race



Source: Henneberger, Cohen, Shipe, & Shaw, 2016

Dual Enrollment by Ethnicity



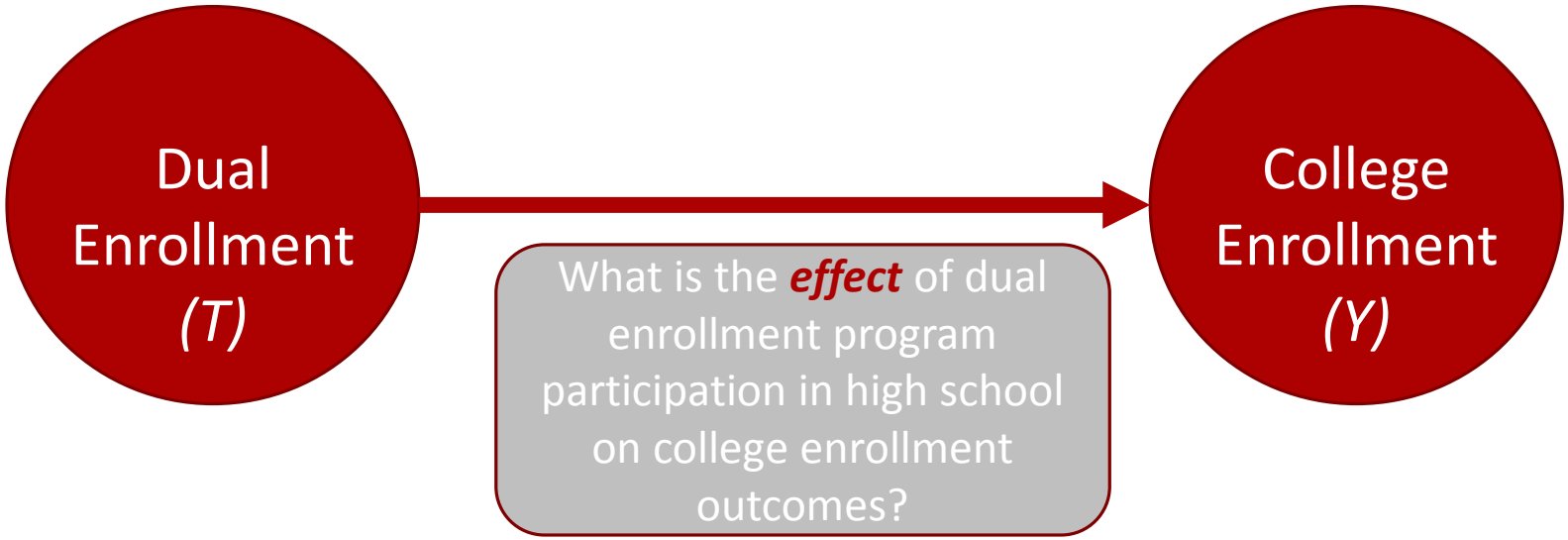
Source: Henneberger, Cohen, Shipe, & Shaw, 2016

Research Question and Motivation

- Motivating Research Question:
  - What is the ***effect*** of dual enrollment program participation in high school on college enrollment outcomes?
  - ***Effect*** implies a causal design where dual enrollment ***causes*** a change in college enrollment.
  - Ideal design = randomization to dual enrollment program and control (Shadish, Cook, & Campbell, 2002).

The Gold Standard: RCT

- Randomize students to participate in dual enrollment or receive no treatment (treatment as usual).
- In this design, each student has a 50% chance to be in treatment group.



MLDS Center
Maryland Longitudinal Data System

State Longitudinal Data and Program Evaluation

- Data compiled in state longitudinal data systems are typically correlational.
- Observations are collected on the same individuals over time without an experimental design.
- No randomization to program and control groups.
  - What is the ***association between*** dual enrollment program participation in high school and college enrollment outcomes?

MLDS Center
Maryland Longitudinal Data System

College Enrollment Outcomes in Maryland

Percentage of Dually Enrolled 12th Grade Students (2013-2014) Who Enrolled in College One Year Later (2014-2015) Compared to the 12th Grade Population

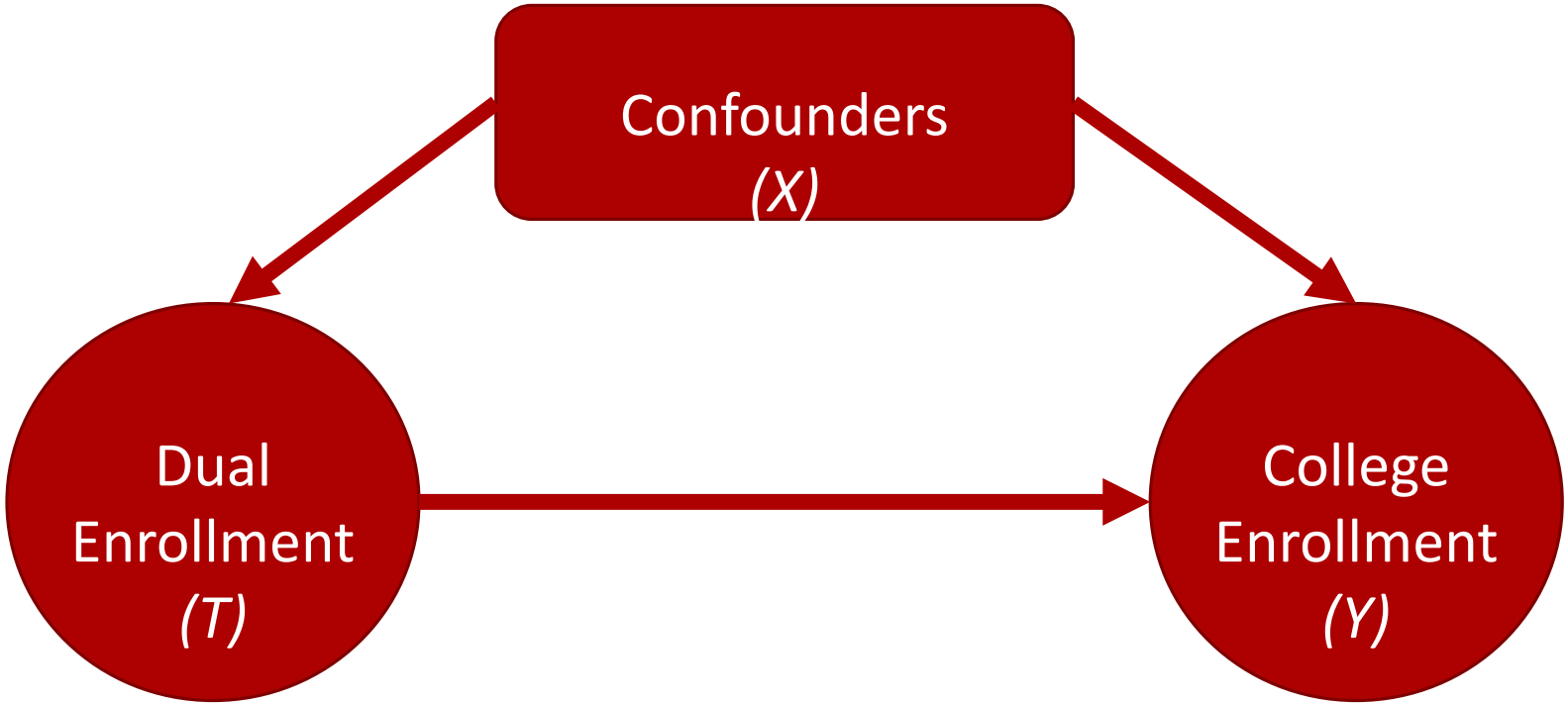
	Dually Enrolled 12th Grade Students (2013-2014)		College Enrollments (2014-2015)	
			All 12th Grade Students	Dually Enrolled 12th Grade Students
	N	%	%	%
Maryland	6,548	11	64	89

Source: Henneberger, Cohen, Shipe, & Shaw, 2016

The Problem: Confounders

- Confounders are variables that are associated with both the treatment (dual enrollment) and the outcome (college enrollment).
- Confounding variables limit the ability to establish causal relationships.
  - Unknown whether any relationship found is due to the treatment or actually due to the confounder.

Correlational Designs

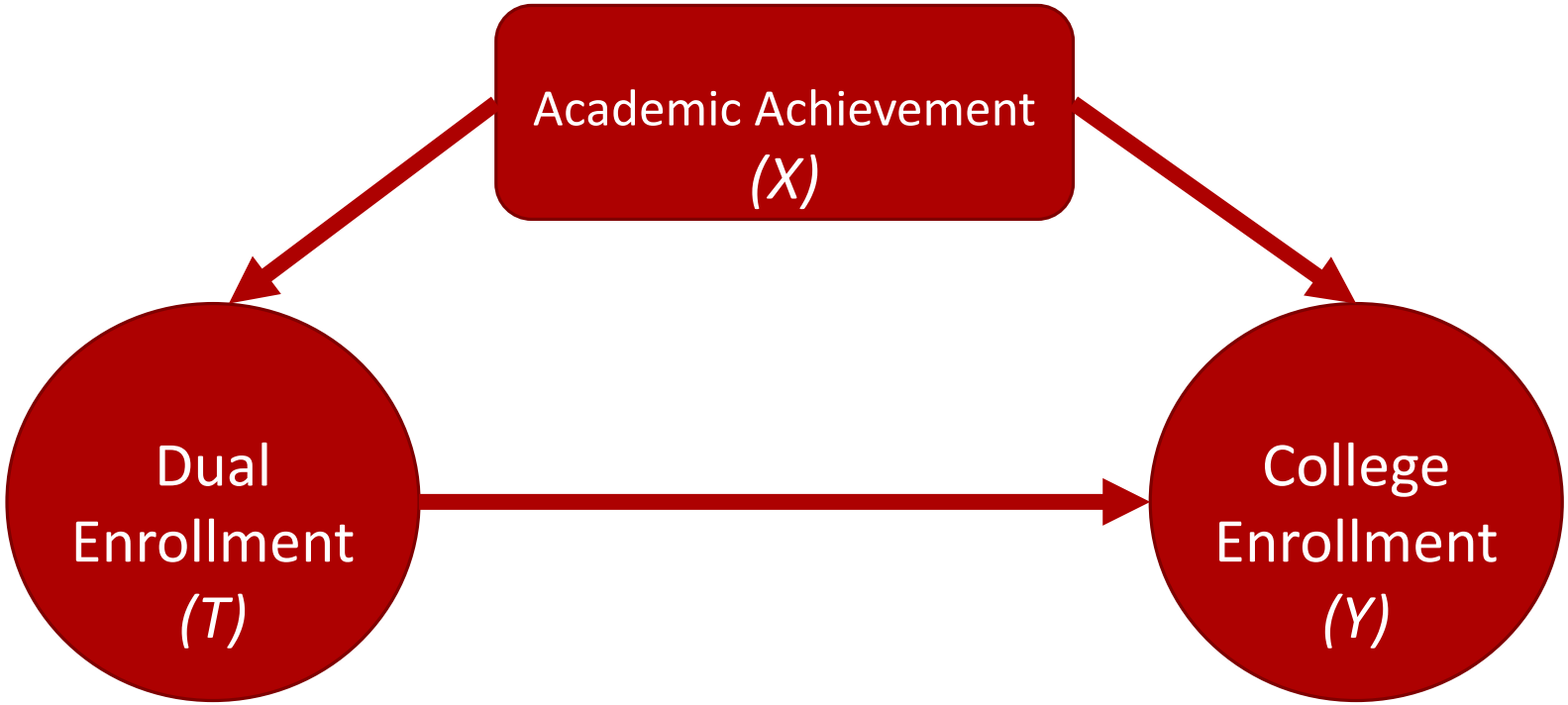


MLDS Center
Maryland Longitudinal Data System

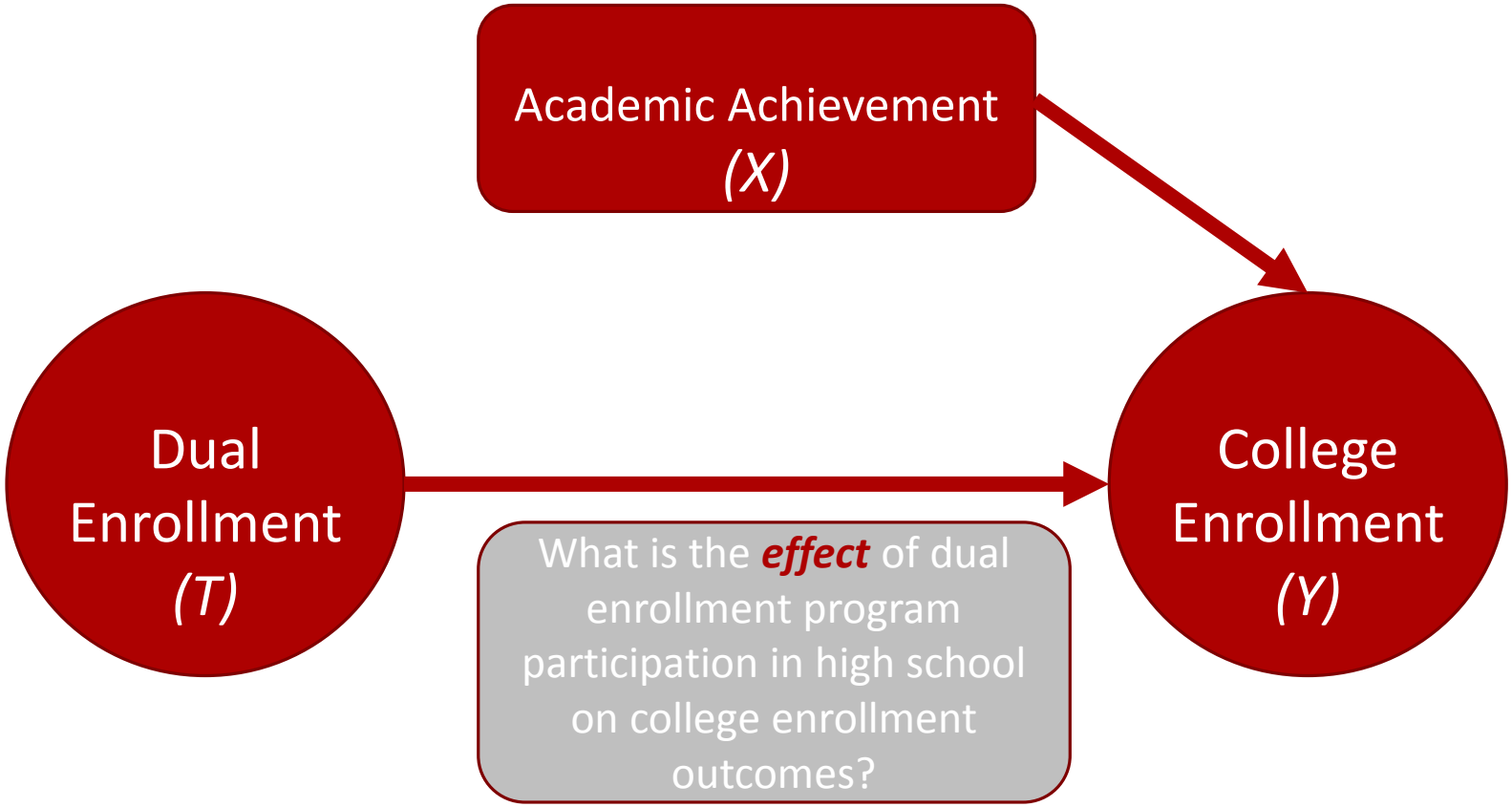
Confounders in Dual Enrollment Program Evaluation

- Gender
- Race/ethnicity
- Socioeconomic status
- High school attendance
- Achievement scores
- Prior academic experience

Academic Achievement as an Example



Academic Achievement as an Example



Modern Causal Inference Techniques

- Modern causal inference techniques can be used to account for the absence of random assignment (Schafer & Kang, 2008).
- Propensity Score Methods
  - Propensity score is the conditional probability of experiencing the treatment given individual's values on confounders (Rosenbaum & Rubin, 1983).
  - The propensity score estimates the probability to participate in the dual enrollment program.
  - Range 0-1; higher = greater likelihood to participate in dual enrollment.
- Improves the ability to make causal inferences about dual enrollment program participation.

MLDS Center
Maryland Longitudinal Data System

Data

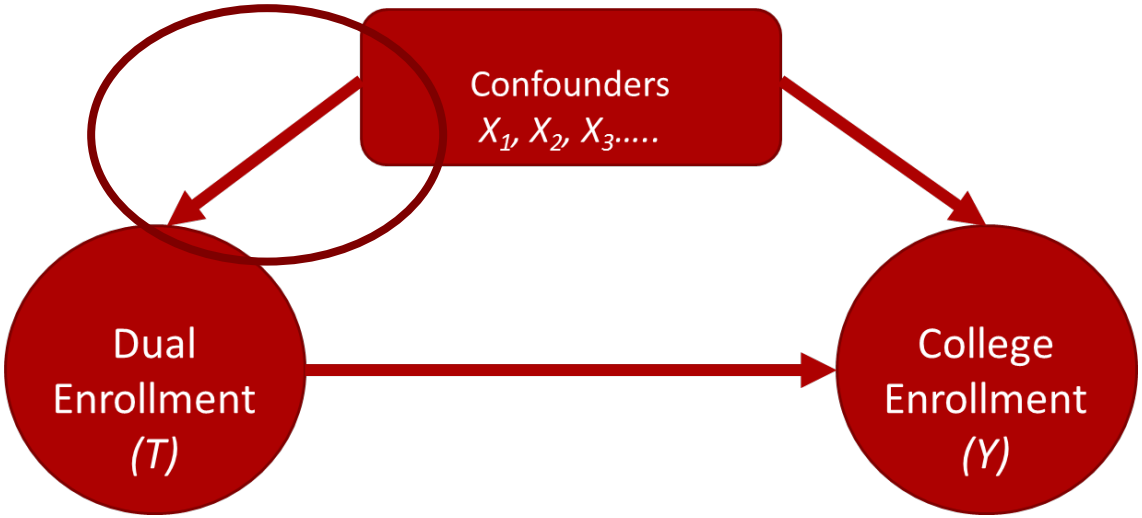
- Data from the Maryland Longitudinal Data System (MLDS)
- Student identified as dually enrolled if:
  - Overlapping enrollment dates in MD public high school and MD college
  - Enrolled in college for at least 30 days
- Population: 62,000 12th grade students (2013-2014)
  - 6,000 of those students were dually enrolled
- Outcome: college enrollment in 2014-2015

Procedure

- Estimate propensity scores
- Create a matched sample
- Check balance using SMD on each covariate
- Check common support
- Outcome model

Estimating Propensity Scores

- Estimated using a logistic regression model
  - Outcome = Dual enrollment (Y/N)
  - Predictors = confounder variables



Results: Estimating Propensity Scores

	Select Logit Coefficients
Male	-0.42***
Black	-0.21***
Hispanic	-0.39***
Free and Reduced Price Meals (FARMs)	-0.32***
Special Education	-0.88***
HSA Algebra	0.01***
HSA English	0.01***
Weeks Absent	-0.10***
Distance to 2-Year College	-0.04***

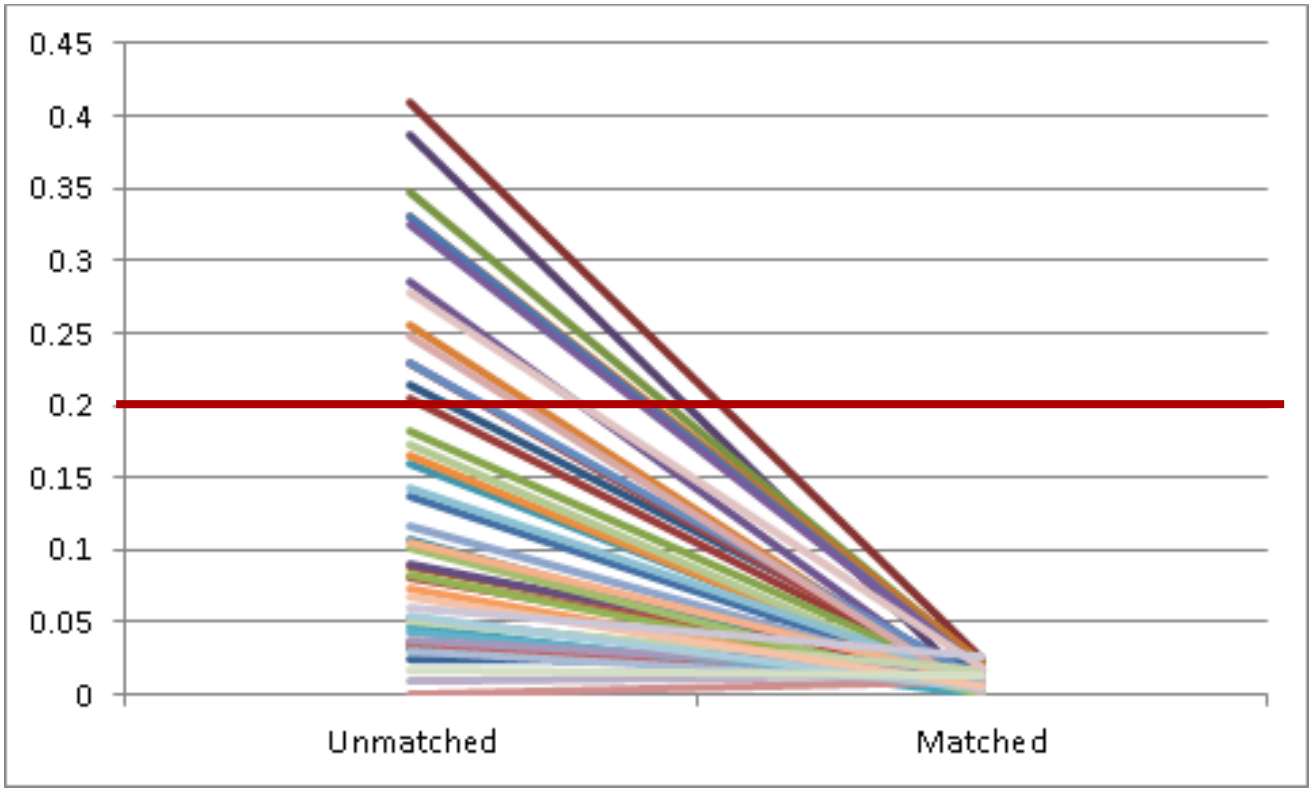
Notes. HSA = high school assessment; additional confounders = homeless; HSA biology; county.
*** $p < .001$

Results: Creating a Matched Sample

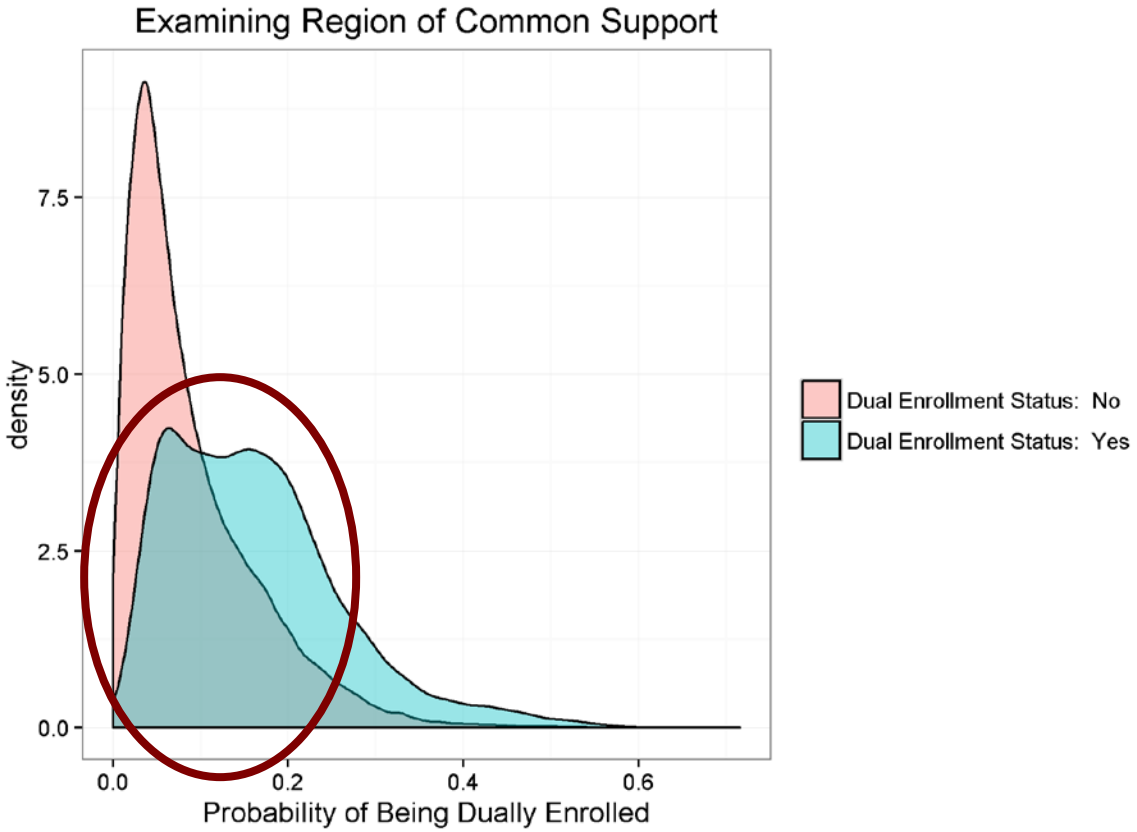
	Control	Treatment (Dual Enrolled)
Total Population	56,000	6,000
Matched	6,000	6,000
Unmatched	50,000	0

Notes. Nearest neighbor matching; 1 to 1 match; caliper = 0.25.

Results: Checking Balance (SMD)

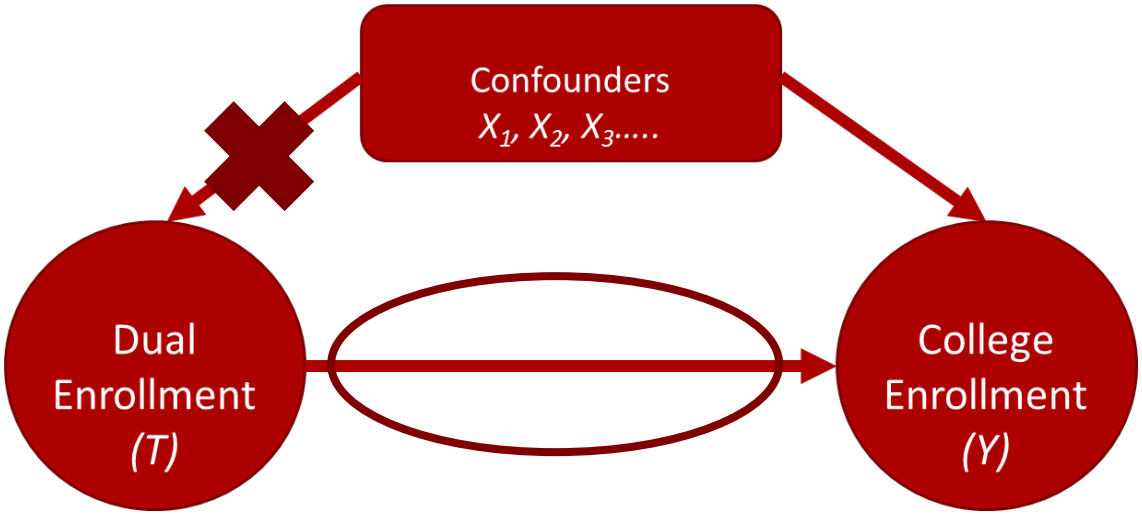


Results: Checking Common Support



Outcome Model

- Estimated using a logistic regression model
  - Outcome = College enrollment
  - Predictor = Dual enrollment (Y/N)



Results: Outcome Model Predicting College Enrollment

	Unmatched (*N* = 62,000) Logit Coefficient	Matched (*N* = 12,000) Logit Coefficient
Dual Enrollment	1.14***	1.13***
	Unmatched Marginal Effect	**Matched Marginal Effect**
Dual Enrollment	0.22	0.16

*** $p < .01$

Interpretation: The predicted probability of enrolling in college is 0.16 greater for a student who was dually enrolled in high school in comparison to a student who was not dually enrolled in high school.

Discussion

- The logit coefficients for dual enrollment predicting college enrollment one year later are of similar magnitude in the unmatched and matched samples.
- The ability to draw causal conclusions about the effect of dual enrollment participation is improved through using **propensity score matching**.
  - This approach gave us the ability to efficiently control for >25 confounding variables.
- **Propensity score matching** is a powerful statistical tool that helps to answer program evaluation questions in the absence of randomization.

For More Information

https://mldscenter.maryland.gov/

Angela K. Henneberger
Research Director
Maryland Longitudinal Data System Center
Angela.henneberger@maryland.gov

MLDS Center
Maryland Longitudinal Data System

Acknowledgement

We are grateful for the data, technical, and research support provided by the MLDS Center and its agency partners. The views and opinions expressed are those of the authors and do not necessarily represent the views of the MLDS Center or its agency partners.

The MLDS Center is an independent agency of the State of Maryland. The mission of the Center is to develop and maintain the Maryland Longitudinal Data System in order to provide analyses, produce relevant information, and inform choices to improve student and workforce outcomes in the State of Maryland.