Estimating the Causal Impact of Tutoring on Student Success Using Machine Learning–Based Propensity Score Matching

Presenter: Max (Zhi) Zhang; Jacqueline Lucero

Institution: Institutional Research and Effectiveness Services

Alamo Colleges District

ALAMO COLLEGES DISTRICT

Malcolm Baldrige National Quality Award
Award Recipient
2024 2018

ACHIEVING THE DREAM
LEADER COLLEGE
OF DISTINCTION

Great Place To Work.
Certified
MAR 2024-MAR 2025
USA

Conceptual Framework & Key Definitions

Tutoring (Intervention):
Tutoring is defined as an institutional academic support intervention delivered outside the classroom and modeled as a binary treatment (received tutoring vs. not).

Student Success (Broad Construct):
Student success is a broad, multidimensional concept encompassing academic progress, persistence, completion, and other indicators of positive student outcomes.

Outcome Measurement in This Study:
In this analysis, student success is measured using Fall-to-Fall persistence, a clear and policy-relevant indicator of continued enrollment.

The Evaluation Challenge:
Because participation in tutoring is non-random, observed differences in persistence cannot be interpreted as causal effects.

Analytic Requirement:
A quasi-experimental framework is required to isolate the effect of tutoring from pre-existing student characteristics.

Methodology & Operationalization

- **Time Frame:** Fall 2024 entry cohort tracked through Fall 2025.
- **Sampling:** Full population used to derive a matched analytic sample of 2,322 students.
- **Outcome Measurement:** Fall-to-Fall persistence measured as a binary indicator (PERSIST_F2F).
- **Baseline Controls:** Rich pre-treatment student characteristics used to capture baseline differences prior to tutoring exposure.

The Multi-Dimensional Control Framework

Demographic Characteristics:
Gender, age, and race/ethnicity.

Neighborhood & Life Context:
High school ZIP code (neighborhood proxy), single-parent status, marital status, and county of residence.

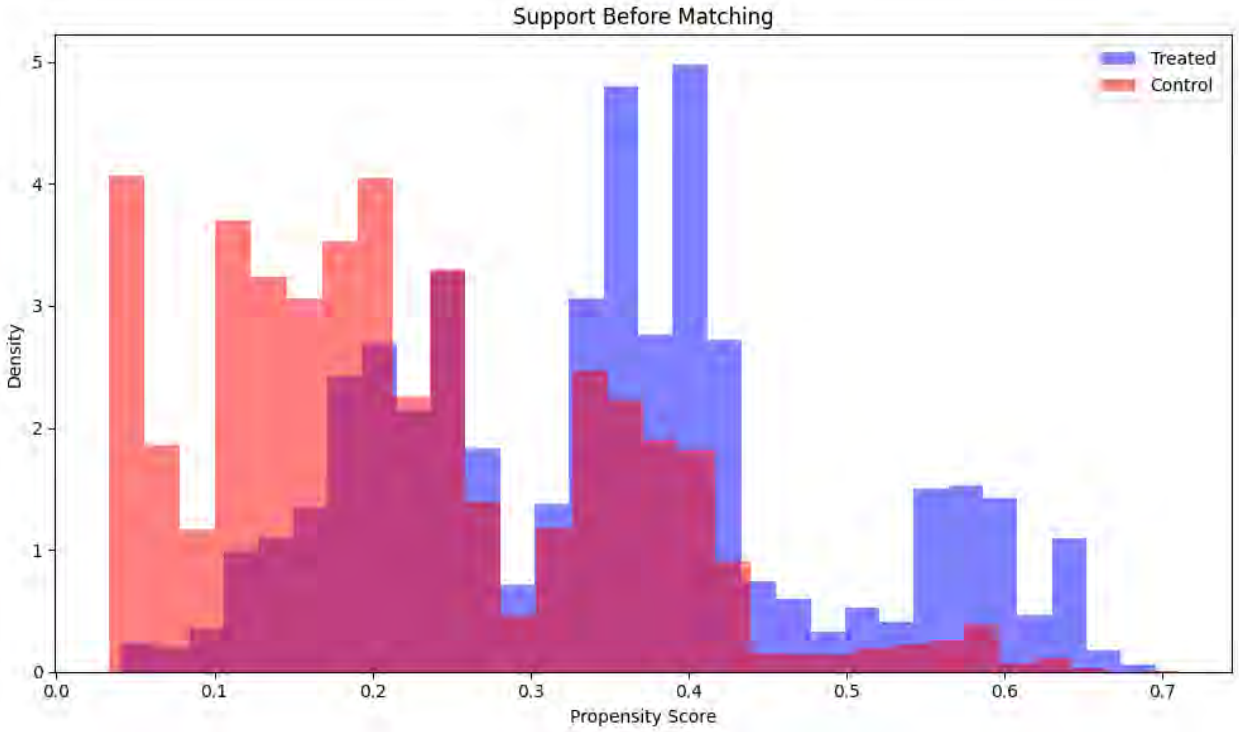
Academic Momentum:
Course load, major type (academic vs. technical), and enrollment intensity (full-time vs. part-time).

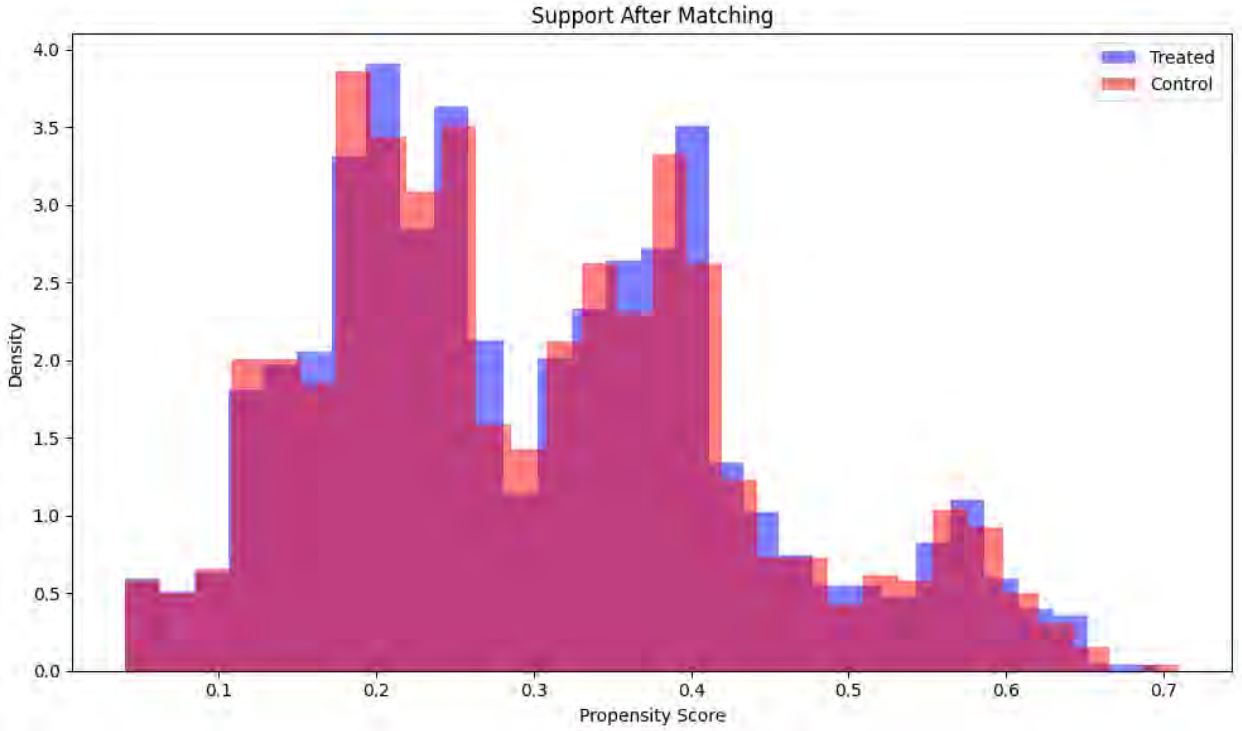
Institutional Context:
College location and tuition status to account for institutional-level differences.

Model Quality & Matching Performance

- **Propensity Model:**
  Gradient Boosting Classifier used to flexibly model non-linear relationships and interactions among baseline covariates.
- **Model Quality:**
  Area Under the Curve (AUC) = 0.731, indicating good separation between tutored and non-tutored students.
- **Matching Criterion:**
  Strict caliper of 0.05 applied to ensure close matches in estimated propensity scores.
- **Matching Outcome:**
  1,161 matched pairs retained (40% of treated students), prioritizing match quality over sample size.

Common Support Before and After Matching





Covariate Balance Diagnostics

Purpose:
To verify that treated and control students are statistically comparable across all observed baseline characteristics.

Balance Assessment:
Standardized Mean Differences (SMDs) were examined for all baseline covariates before and after matching.

Key Result:
After matching, SMDs across all covariates were reduced to near zero and well below conventional imbalance thresholds.

Interpretation:
The matched groups are statistically indistinguishable across observed characteristics, isolating tutoring as the primary remaining difference.

Core Findings: The 17.1% Persistence Lift

Matched Sample Results (n = 2,322):
Tutored students: **72.5%** Fall-to-Fall persistence
Matched control students: **55.4%** Fall-to-Fall persistence

Estimated Causal Effect:
+17.1 percentage point increase in persistence

Relative Difference:
Tutored students were **30.8% more likely to persist** than their matched peers

Limitations & Scope of Inference

Local Scope of the Estimated Effect:
The estimated impact applies to students within the matched sample—those similar to current tutoring participants and within the region of common support.

Generalizability Considerations:
Students in the matched sample may be more predisposed to academic success than the broader student population, as reflected in baseline indicators such as prior academic performance.

Unobserved Factors:
Although extensive baseline controls were used, unmeasured characteristics (e.g., motivation or help-seeking orientation) may still influence both tutoring participation and outcomes.

Outcome Focus:
Student success is measured using Fall-to-Fall persistence; other dimensions of success remain important but are beyond the scope of this analysis.

Institutional Implications & Advising

Identifying Actionable Risk:
Among students comparable to tutoring participants, those who did not receive tutoring persisted at a substantially lower rate (55.4%), identifying a group currently underutilizing support services.

Proactive Advising Strategy:
The matching framework can be used to proactively identify students who resemble high-risk matched controls and prioritize early outreach and structured connection to tutoring.

Institutional Takeaway:
Tutoring emerges as a high-impact lever for improving persistence among students who are otherwise positioned to succeed.

Thank you.

Learn more at

alamo.edu

ALAMO COLLEGES DISTRICT

Malcolm Baldrige National Quality Award
Award Recipient
2024 2018

ACHIEVING THE DREAM
LEADER COLLEGE
OF DISTINCTION

Great Place To Work.
Certified
MAR 2024-MAR 2025
USA

Scan the QR code to complete the session survey.

2026