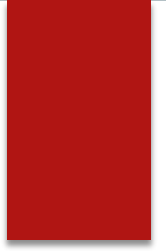




A Friendly Intervention

PROPENSITY SCORES, OVERLAP WEIGHTS, AND PEER ASSISTED LEARNING

The Research Problem



Peer Assisted Learning

- ▶ PAL is a one-unit supplemental instruction course that students can take alongside a number of challenging NSM courses.
- ▶ We know that PAL works!
 - ▶ PAL improves student grades for the associated course by an average of 0.42 points (on a GPA scale).

Does Dosing Matter?

Research question: is the number of PAL courses taken (at any point) predictive of student success?

- ▶ Outcome: course grade
- ▶ Treatment: number of PAL courses taken

The data: 74,544 student-course instances which include grade, PAL participation, general course information, and student demographic information.

Does Dosing Matter?

- ▶ We might think to regress course grade on number of PAL courses.
- ▶ But there are some problems with this approach!
- ▶ It turns out our research question (and our data) are more complex than they sound!

Challenge: Missing Data

- ▶ We remove variables with more than 10% missing.
- ▶ After those are removed, we subset the data to *complete cases*.
- ▶ But we'd like to keep information like high school GPA.
 - ▶ This is missing for most transfer students, so complete cases *bias* the data toward freshmen.
 - ▶ So we create a variable that indicates when HS GPA is missing, and then *impute* (fill in) the missing GPAs with the mean GPA from the non-missing data.

Challenge: Student Self-Selection into PAL

- ▶ Treatment (PAL) assignment is not random, so our data is non-experimental.
 - ▶ Without additional work, we can't say that PAL enrollment *causes* GPA outcomes.
- ▶ Solution: by re-weighting the data, we can compare PAL students to similar students in the non-PAL group.
 - ▶ This way we can "compare like with like" and draw causal conclusions.

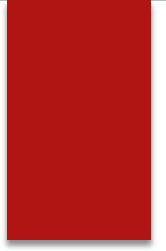
Challenge: Repeated Measures

- ▶ Students appear multiple times in the data.
- ▶ Grades for a single student will correlate with each other, which violates modeling assumptions.
 - ▶ We will need to account for this in our models.

Challenge: Other Factors

- ▶ There are a lot of factors (other than PAL) that impact student grades.
- ▶ Some of these will be averaged out when we match similar students between PAL and non-PAL, but it's not practical to tackle all of these in the matching.
 - ▶ For example, matching on course instructor is not feasible.
- ▶ Solution: include these factors in our models.
 - ▶ If PAL dosage is still useful, even when these factors are accounted for, that strengthens our conclusions.

Overview of Methodology



Propensity Score Weighting

- ▶ A *propensity score* is a student's probability of enrolling in PAL, given their academic baseline and demographic information.
 - ▶ We can use propensity scores to match similar PAL and non-PAL students.
 - ▶ We can also use these to calculate **weights** to use in future models, which accomplishes essentially the same thing as matching.

Propensity Score Weighting

- ▶ A gradient boosting machine (a statistical learning technique) was used to calculate the weights.
- ▶ We examine the *average treatment effect in the overlap* (ATO).
- ▶ ATO assigns (overlap) weights to each student based on their probability of PAL participation.
 - ▶ Students who would definitely take PAL or definitely avoid it get very little weight.
 - ▶ Students who could plausibly go either way are weighted most heavily.

Repeated Measures and Other Factors

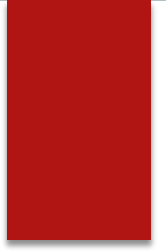
The model itself is basically a linear regression, but with some additional complexity:

- ▶ Control for repeated measures by treating student as a *cluster*.
- ▶ Control for other factors by including them as main effects in the model.
 - ▶ Instructor, course, major.
- ▶ The overlap weights also go into the model call!

Contrasts

- ▶ *Contrasts* allow us to perform pairwise comparisons between different treatment levels.
 - ▶ We take into account multiple comparisons, which can increase the chance of false positive.
- ▶ In general, we are interested in whether each additional PAL taken increases grade beyond the previous level.
 - ▶ That is, we want to compare 0 to 1, 1 to 2, and 2 to 3+ PAL courses.

Results



Dosing Works!

Contrast	Estimate	P-Value
1 - 0	0.122	<0.0001
2 - 1	0.113	0.0002
(3+) - 2	0.106	0.0083

- ▶ The *estimate* is the mean increase in grade with the additional PAL course.
 - ▶ That is, students who've taken 2 PAL courses see grades 0.113 points higher than those who've taken only 1 PAL course.

Decoupling of Prior GPA

An interaction model between PAL and baseline GPA shows:

Dosage	GPA Trend	95% Confidence Interval
0	0.862	(0.807, 0.918)
1	0.713	(0.640, 0.786)
2	0.660	(0.555, 0.765)
3+	0.590	(0.477, 0.702)

- ▶ In general, baseline GPA is highly predictive of future grades.
- ▶ But as PAL participation increases, prior GPA matters less and less
 - ▶ That is, PAL helps level the playing field!

URM vs Non-URM students

URM Students

Contrast	Estimate	P-Value
1 - 0	0.1667	<0.0001
2 - 1	0.1294	0.0026
(3+) - 2	0.1320	0.0404

Non-URM Students

Contrast	Estimate	P-Value
1 - 0	0.1442	<0.0001
2 - 1	0.1324	0.0045
(3+) - 2	0.0893	0.3088

- PAL benefits both groups, but we see a plateau for non-URM students at higher doses.
- Non-significant (large) p-values mean we cannot conclusively tell if our estimates are nonzero.

Native Freshmen vs Transfer

Native Freshmen

Contrast	Estimate	P-Value
1 - 0	0.1562	<0.0001
2 - 1	0.1147	0.0012
(3+) - 2	0.1478	0.0040

Transfer

Contrast	Estimate	P-Value
1 - 0	0.1537	0.0002
2 - 1	0.1606	0.0094
(3+) - 2	-0.0217	0.9891

- Additional PAL benefits native freshmen at all levels.
- High dose PAL benefits are less consistent for transfer students.
 - Transfer students arrive with different skills, take different courses.

Sensitivity Analysis

We might wonder:

- ▶ What about the stuff we didn't measure?
- ▶ Might that explain away the effect of PAL?

A sensitivity analysis and E-values allow us to quantify:

- ▶ How strong would an unmeasured factor need to be in order to explain away our results?

Sensitivity Analysis

Comparison	Estimate	E Value
1 - 0	0.122	1.404
2 - 0	0.235	1.645
(3+) - 0	0.340	1.864

- ▶ To fully explain away the 3+ PAL course result, an unmeasured factor would need to have a very strong relationship with both PAL participation and grades.
 - ▶ This is unlikely (and that's good!)

Questions?

