

Using Enrollment Data to Predict Retention Rate

*Natalie Walleser Solverson, MPA
Melissa A. Bingham, PhD
University of Wisconsin-La Crosse*

Retention...

Was it the



or the



?

Why do we care?

- Student Success
- This Thing
- College is Still Worth It

Guiding Research Question

How “good” is the UW-L retention rate of first-time, full-time students to the second year?

...as in, how much of the retention rate is student characteristics, and how much is university intervention?

What we'll cover today

- What we did and a little about why we did it
 - predicted versus observed retention
- What challenges we faced
- What to consider going forward

A little about UWL

- selective, 10,000 student public university
- 91 undergraduate programs
- biology most popular major; draws students interested in health careers

Retention Rates

Group	Retention Rate
UWL FTFT Cohort Fall 2009	85.5%
UWL FTFT Cohort Fall 2010	84.9%
UWL FTFT Cohort Fall 2011	85.9%
Comparisons	
ACT comparison group	83.9%
UW System Comprehensives	76.8%

Literature Review

- Test scores and high school GPA exhibit a strong positive relationship with retention (Porter and Tinto as cited in Delen, 2011-2012).
- Integration to campus community cited as strong influence in retention (Tinto, seminal work in 1975).
- “Despite the extensive body of literature devoted to studying student persistence, overall retention rates and graduation rates have not improved over time” (Seidman, 2012).

What else is out there?

- Predictive modeling at the individual student level
- Astin (1997) and Woodard, Mallory, and De Luca (2001) modeled graduation rates
- Could not find predicted-versus-observed retention rate as a dependent variable

Data Collection

- Collected from UW System Central Data Request (CDR)
 - gender
 - residency status (WI/MN/other)
 - ACT composite score
 - high school class rank (as a percentage)
 - race/ethnicity - international students not considered
 - first generation student?
 - math remediation?
 - total need based financial aid
- Variable limitations because of what is available from CDR and student privacy issues and new data collection policy

What we didn't have

- Housing
- Employment
- Major (declared versus undeclared)
- others?

Data Collection

	UW-L (n=5448)	Other Comprehensives (n=30175)
Median ACT Score	25	22
Median HS Class Rank	82	68
% First Generation	37.6%	53.2%
% Placed in Remedial Math	3.5%	21.1%
Avg Need Based Aid Financial Aid Package	2.828	3.858

UW-L students are have higher academic achievement in high school and demonstrate less financial need than students at the other comprehensive campuses (as a whole).

Modeling Retention Rate

- Used logistic regression to model first to second year retention (retained=1, not retained=0) dependent on student inputs.
- Used data from the UW Comprehensives, excluding UWL.
- All variables but gender (male/female) and residency (WI/MN/other) were found to be significant predictors of retention.

Variables Used in the Final Model

Variable	Description
ACT	ACT composite score
Rank	High school class rank as a percentage
SEAsian	Coded as 1 if Southeast Asian American, 0 otherwise
White	Coded as 1 if White/Caucasian, 0 otherwise
FirstGen	Coded as 1 if a first generation college student, 0 otherwise
Need	Total need-based financial aid, in thousands of dollars
RemMath	Coded as 1 if required to take a remedial math course, 0 otherwise

Logistic Regression Coefficients

	Coefficient*	exp(Coefficient)	p-value
Intercept	-1.0506	0.350	0.000
ACT	0.0190	1.019	0.000
Rank	0.0235	1.024	0.000
SEAsian	0.2201	1.246	0.043
White	0.3604	1.434	0.000
FirstGen	-0.2720	0.762	0.000
Need	0.0424	1.043	0.000
RemMath	-0.1611	0.851	0.000

*Coefficients for modeling the log odds of being retained

Interpretation of Coefficients

- Positive relationship with retention:
 - Higher ACT Score
 - Higher Class Rank
 - More Need-Based Financial Aid
- First Generation = 76.2% odds of being retained compared to either/both parents having bachelor degree
- White students most likely to be retained
 - Southeast Asian 24.6% odds of being retained compared to all other students of color

Comparing Predicted and Actual Retention Rates (Overall)

- UW-L averages (ACT=24.768, Rank=80.729, Need=2.828, FirstGen=.3758, RemMath=.0352, SEAsian=.0174, White=.9214) were plugged into

$$\ln(p/(1-p)) = -1.0506 + 0.0190 * \text{ACT} + 0.0235 * \text{Rank} + 0.2201 * \text{SEAsian} + 0.3604 * \text{White} - 0.2720 * \text{FirstGen} \\ + 0.0424 * \text{Need} - 0.1611 * \text{RemMath}$$

- Log odds of 1.665 resulted, which converts to a 84.1% predicted retention rate.
- This is lower than UW-L's actual overall 85.7% rate.

Comparing Predicted and Actual Retention Rates (Subgroups)

- Using the summary statistics on the previous slide with appropriate 0's and 1's, we can find predicted retention rates for the subgroups.
- For example, when considering SE Asian students who are first generation, we entered into the equation $\text{White}=0$, $\text{SEAsian}=1$, $\text{FirstGen}=1$, $\text{ACT}=19.442$, $\text{Rank}=68.302$, $\text{Need}=11.539$, $\text{RemMath}=.1512$.

Comparing Predicted and Actual Retention Rates (Subgroups)

- Because of the large number of STEM majors at UW-L, the ability to progress in math is particularly important.
- For students taking remedial math (RemMath=1), we have averages ACT=21, Rank=71, Need=5.1, White=.708, SEAsian=.078, FirstGen=.495.
- These values yield a predicted retention probability of 77.0%, while the actual UW-L retention rate for this subgroup is 82.8%.

UW-L Summary Stats by Subgroups

First-Generation Status	Race/Ethnicity (U.S. students)	Average ACT	Average Rank	Average Need in Thousands	% needing Remedial Math
Neither Parent w/ Bachelor's	White	24.606	82.429	3.963	3.12
	Southeast Asian	19.442	68.302	11.53	15.12
	Other	22.057	69.483	7.066	17.86
At Least One Parent w/ Bachelor's	White	25.203	81.241	1.723	2.47
	Southeast Asian	20.222	69.444	9.026	22.22
	Other	23.758	69.813	3.075	8.25

Comparing Predicted and Actual Retention Rates (Subgroups)

First Generation Status	Race/Ethnicity (U.S. students)	Predicted Probability of Retention	Actual Retention Rate	Difference (Actual - Predicted)
Neither Parent w/ Bachelor's	White	83.32%	85.32% (n=1826)	2.00%
	Southeast Asian	79.25%	83.72% (n=86)	4.47%
	Other	73.17%	77.86% (n=140)	4.69%
At Least One Parent w/ Bachelor's	White	85.45%	86.47% (n=3193)	1.02%
	Southeast Asian	82.28%	88.89% (n=9)	6.61%
	Other	76.16%	83.51% (n=194)	7.35%

UW-L students have higher retention rates than what is predicted across all subgroups

Other things we did

- The model was considered “adequate”
- What else did we do to try to improve:
 - Subset the data into those students from the other comprehensives who would likely have been admitted to UWL
 - Adding interaction terms
- These efforts did not make a huge difference
 - missing variables

Limitations/Improvements

- UW System comprehensives are not all the same
 - Perhaps a different combination of institutions
- Use own data
 - Build out from previous years, apply to subsequent years

Discussion

- The overall actual retention rate at UW-L (85.7%) is higher than what would be seen at other UW Comprehensives (84.1%) given the same student inputs.
- Large differences found in the predicted and actual retention rates for students of color, providing additional evidence that the supports and services provided to these students are positively impacting the retention rates.

What might be helping?

- On campus
 - Strong residence life (97% first years live on)
 - Attention to Campus Climate
- Support for Underresourced Students
 - Starting to see uptick in retention of low-income students of color in last couple years
- MOOC and relatively low remediation levels
- Relative affluence

What to do with information?

- Monitor and respond
 - What to do if retention rate dips?
- Set appropriate goals
 - If you have an achievement gap - how big and what to do to close it?