LA-STEM Student Outcome Report

Tim Weston

September 4, 2009

Executive Summary

The analysis examined if LA-STEM students experience beneficial outcomes related to graduation, GPA and attaining STEM degrees when compared to students from the general LSU population entering the university with an intended STEM major.

Analysis used propensity score logistic regression to adjust for preexisting differences in student characteristics between LA-STEM and non-LA-STEM groups.

The student characteristics of racial/ethnic status (being African-American or not being White), being an instate student, attending public school, ACT math score and HS GPA all predicted if a student was a member of LA-STEM.

Comparisons of outcomes showed:

Students in LA-STEM who started in either 2004 or 2005 cohorts graduated in higher rates than their peers in the general population.

Students withdrawing from LA-STEM had a graduation rate statistically equivalent to students in the general population. Both of these groups graduated at lower rates than the LA-STEM group.

The LA-STEM cohort group earned a higher cumulative GPA than the general population, even when the propensity matching score variable was held constant.

Students staying in LA-STEM had a significantly higher GPA than students who withdrew from LA-STEM, and students from the general population.

Students withdrawing from LA-STEM showed a significantly lower average number of credit hours taken per semester than students in the LA-STEM and general population groups.

All LA-STEM students graduated with STEM degrees. Only one student from the group of LA-STEM students withdrawing from the program graduated with a non-STEM degree, while 26% of students in the general population switched from their intended major out of STEM.

Overall, students who withdrew from the LA-STEM program were more likely to be African-American, and less likely to be white than their peers in both groups. Their HS GPA was slightly lower, and they scored lower on ACT comp and ACT math tests than those who did not withdraw.

Introduction

This analysis examined if LA-STEM students experienced beneficial outcomes related to graduation, GPA and attaining STEM degrees when compared to students from the general LSU population entering the university with an intended STEM major.

Using quantitative variables, participating students were matched with non-participating students using propensity score logistic regression. Student characteristics such as high school grade GPA, ACT scores, and parental income were used to predict group membership (LA-STEM, or non-LA-STEM). The values for predicted group membership were then incorporated in statistical models to hold preexisting differences constant between groups, and thus make a fairer comparison of outcome variables such as graduation rates and university level GPA.

Data Preparation and Considerations

The data consisted of 20,148 records from 3,488 students; each student had an individual record for each semester enrolled. Some variables were constant across semesters, while others (such as GPA) varied across semesters. Variables were aggregated to give each student only one record representing all semesters enrolled. To aggregate constant variables, the first record in the series was used. Variables that changed across semesters were aggregated using the series average. For the graduation variable, the last record (indicating graduation) was used in the aggregate data file.

Some variables needed to be recoded so that they could be used in statistical procedures. The racial/ethnic variable was recoded into a series of dichotomous variables such that African-American (and each other group) were represented by a 1, and all other groups represented by a 0. Some textual variables (e.g. *Gender*) were also converted into numerical values so that they could be used in quantitative analysis.

Group variables

LA-STEM students were classified as entering the program (based on *lastemi* = student in LA-STEM program in entry term (y/n)), and as either staying in the program through the last semester of data collected, or withdrawing from the program (based on *lastem* = LA-Stem (y/n) by semester).

The first LA-STEM Group (called *LA-STEM COHORT GROUP* in this analysis), consisted of 133 students in the LA-STEM group. The second group (called *LA-STEM PARTICIPATION GROUP*) split the first group, with 53 students classified as withdrawing from LA-STEM, and 80 students staying in the program.

Propensity Analysis and Summary Data

Propensity analysis used Logistic Regression to match groups on a series of variables characterizing the comparison groups. The logistic regression equation calculates the probability of group membership (in this case being a *LA-STEM COHORT GROUP* participant) given characteristics such as racial/ethnic status, family income or academic achievement. A specific probability of group membership was then used to make a fair comparison by matching LA-STEM and non-LASTEM students on the propensity variable and comparing outcomes between groups using Analysis of Covariance (ANCOVA).

The logistic regression formula used in this analysis is:

Group membership probability (LA-STEM COHORT GROUP =1) = $1/(1 + e^{-z})$

$z = \beta_1 X_1 + \beta_2 X_2 \ldots\ldots \beta_m X_m + C$

Where β are logistic regression weights calculated to maximize predictive power, C is an intercept constant and X are values of independent variables.

Variables were entered into the regression equation and retained if they were statistically significant predictors at the $p < .10$ level. The variables retained and their means, standard deviations and N for each group are shown in Table 1. Variables attempted, but not entered into the equation are listed in table 2. The actual propensity formula including beta (β) weights, WALD statistics and p-values is presented in table 3.

	LASTEM COHORT GROUP					
	NON-LA STEM			LA STEM		
	Mean	Standard Deviation	Valid N	Mean	Standard Deviation	Valid N
AFRICAN-AMERICAN	.04	.21	3355	.38	.49	133
WHITE	.85	.35	3355	.47	.50	133
INSTATE	.88	.32	3355	.92	.26	133
PUBLIC SCHOOL	.59	.49	3355	.77	.42	133
ACT MATH	27.24	3.48	3262	28.44	3.88	129
HS GPA	3.86	.15	3251	3.84	.17	131

All decimals represent proportions

Table 1 Means, Standard Deviations and N for LA STEM and Non-LA STEM students.

	LASTEM COHORT GROUP					
	NON-LA STEM			LA STEM		
	Mean	Standard Deviation	Valid N	Mean	Standard Deviation	Valid N
AGE	21.23	1.53	3355	21.34	1.56	133
GENDER (proportion male)	.52	.50	3355	.51	.50	133
HISPANIC	.02	.14	3355	.05	.22	133
ASIAN	.06	.23	3355	.07	.25	133
ACT COMP	27.71	2.74	3262	28.22	3.05	129
ACT ENGLISH	28.44	3.69	3262	29.14	3.43	129
HS GPA (LSU CALC)	3.80	.16	3355	3.77	.24	132
HS ENROLLMENT	3.34	1.13	2987	3.13	1.11	126
FAMILY INCOME	127448.29	116431.89	3079	86614.67	45704.01	123
HS RANK	21.06	26.55	3014	24.72	29.11	127

Table 2 Means, SD and N for variables not entered in propensity analysis.

	β	S.E.	Wald	df	Sig.	Exp(B)
AFRICAN-AMERICAN	2.138	.312	46.954	1	.000	8.483
WHITE	-.976	.272	12.864	1	.000	.377
INSTATE	.955	.419	5.191	1	.023	2.598
PUBLIC SCHOOL	.416	.226	3.377	1	.066	1.516
ACT MATH	.179	.028	41.730	1	.000	1.196
HS GPA	-1.235	.672	3.378	1	.066	.291
Constant	-4.307	2.664	2.615	1	.106	.013

Table 3 Propensity Analysis Logistic Regression equation for LA-STEM analysis.

Data considerations for propensity analysis

Because of the very large disparity in the size of the groups (causing restricted variability in the dependent measure), a "bootstrap" procedure was conducted comparing 25 smaller (n =1000) random samples of the larger group (Non-LA-STEM) with the LA-STEM group. No large differences in the composition or numerical values of beta weights were found between the smaller samples and the sample as a whole, moreover, using group probability weights derived from the smaller samples did not substantially change how the propensity variable worked as a matching covariate in outcome comparisons. Therefore, the original formula from the whole sample was retained.

Some variables also entered the formula but were not used because their amount of missing data limited the number of cases allowed in outcome analyses. For instance, *Family Income*, while predictive of group membership, could not be used in the logistic regression procedure because of the amount of missing data.

Outcome variables

Several outcome variables were used to compare LA-STEM and non-LA-STEM groups. These variables included *Graduation Rate, Cumulative GPA, Average Credit Hours per Semester, Average Number of Withdraws* and match between intended major and actual degree for STEM majors.

Graduation rate was calculated only for those students who had stayed in school for four years or more (e.g., starting before Spring 2006) and thus had the time to graduate; all students who graduated were counted, while only those who had been in school long enough to potentially graduate were counted as not receiving a degree. Forty-six LA-STEM students received degrees.

The Cumulative Grade Point Average is the average of all semesters' grades for each student. *Average Credit Hours per Semester* was calculated from non-zero records of credits taken each semester, and *Average Number of Withdraws* is the average number of courses with a "w" (withdrawal) grade over all semesters enrolled. Mean, SD and N for each group are presented in table 4 for all outcome variables for the cohort group, and table 5 for the participation group.

	LASTEM COHORT GROUP					
	NON-LA STEM			LA STEM		
	Mean	Standard Deviation	Valid N	Mean	Standard Deviation	Valid N
GRADUATION RATE	.54	.50	1271	.77	.43	60
CUMULATIVE GPA	3.24	.63	3337	3.46	.43	133
AVERAGE SEMESTER HOURS	14.33	1.77	3347	13.73	2.09	133
AVERAGE NUMBER OF WITHDRAWS	.37	.56	3347	.28	.33	133

Table 4 Mean, SD and N for outcome variables for LA-STEM Cohort Groups

	LA STEM PARTICIPATION GROUP								
	NON-LASTEM			WITHDREW LA STEM			LA STEM		
	Mean	Standard Deviation	Valid N	Mean	Standard Deviation	Valid N	Mean	Standard Deviation	Valid N
GRADUATION RATE	.54	.50	1271	.58	.50	31	.97	.19	29
CUMULATIVE GPA	3.24	.63	3337	3.27	.48	53	3.58	.36	80
AVERAGE SEMESTER HOURS	14.33	1.77	3347	13.10	2.35	53	14.15	1.79	80
AVERAGE NUMBER OF WITHDRAWS	.37	.56	3347	.40	.38	53	.20	.28	80

Table 5 Mean, SD and N for outcome variables for LA-STEM Participation Groups

RESULTS

1. *Do LA-STEM program participants graduate at a higher rate than students in the general population?*

The analysis first compares LA-STEM and non-LA-STEM students who either started school before Spring 2006 and did not receive a degree, or who received a degree. Table 7 shows the cross-tabulation of group and frequency of degrees with both actual and expected frequencies. LA-STEM students graduated at a higher frequency than expected. The result shows a significant Chi-Square of $\chi^2 = 12.0$, df = 1, p = .001.

			DEGREE GRANTED		
			No	Yes	Total
LASTEMCODE	NON-LA STEM	Count	586	685	1271
		Expected Count	573.0	698.0	1271.0
	LA STEM	Count	14	46	60
		Expected Count	27.0	33.0	60.0
Total		Count	600	731	1331
		Expected Count	600.0	731.0	1331.0

Table 6 Cross-tabulation for LASTEM COHORT and DEGREE GRANTED

Using graduation rate as a proportion in the Analysis of Covariance with the propensity score matching variable as a covariate showed a similar result as the chi-square with F = 10.02, df 1, 1192, p = .002. The propensity matching variable had little effect on the outcome of the analysis (F = 1.12, p =.288).

Dependent Variable: DEGREEGRANTED

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	3.701[a]	2	1.851	7.521	.001
Intercept	63.178	1	63.178	256.758	.000
MATCHING	.278	1	.278	1.129	.288
LASTEM COHORT	2.474	1	2.474	10.054	.002
Error	292.813	1190	.246		
Total	642.000	1193			
Corrected Total	296.515	1192			

a. R Squared = .012 (Adjusted R Squared = .011)

Table 7 ANCOVA table for DEGREE GRANTED comparison, LA-STEM COHORT Group

When the LASTEM PARTICIPATION GROUP variable (including students withdrawing from LA-STEM) was used in the analysis, the Chi-Square was also significant at $\chi^2 = 20.9$, $p < .0001$. The pattern of frequencies shows much higher than expected frequencies for the LA-STEM group.

The ANCOVA comparison for the same group shows a significant main effect at $F = 8.6$, $df = 2,1191$, $p < .0001$. A post-hoc test showed no significant differences between the WITHDREW LA-STEM and general population on this variable, but did show significant differences between the LA-STEM group and both the general student population and WITHDREW LA-STEM group.

			DEGREE GRANTED		
			No	Yes	Total
LA STEM PARTICIPATION GROUP	NON-LASTEM	Count	586	685	1271
		Expected Count	573.0	698.0	1271.0
	WITHDREW LA STEM	Count	13	18	31
		Expected Count	14.0	17.0	31.0
	LA STEM	Count	1	28	29
		Expected Count	13.1	15.9	29.0
Total		Count	600	731	1331
		Expected Count	600.0	731.0	1331.0

Table 8 Cross-tabulation for LASTEM PARTICIPATION and DEGREE GRANTED

Dependent Variable: DEGREE GRANTED

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	5.442[a]	3	1.814	7.410	.000
Intercept	43.362	1	43.362	177.130	.000
MATCHING	.278	1	.278	1.135	.287
LA STEM PARTICIPATION	4.214	2	2.107	8.607	.000
Error	291.073	1189	.245		
Total	642.000	1193			
Corrected Total	296.515	1192			

a. R Squared = .018 (Adjusted R Squared = .016)

Table 9 ANCOVA table for DEGREE GRANTED comparison, LA-STEM PARTICIPATION Group

The analysis shows that students in the LA-STEM who started in either 2004 or 2005 cohorts graduated in higher rates than their peers in the general population.

Students withdrawing from LA-STEM had a graduation rate statistically equivalent to students in the general population. Both of these groups graduated at lower rates than the LA-STEM group.

2. Do LA-STEM program participants have higher cumulative grade point average?

Students' university cumulative grade point average was compared between LA-STEM and non-LA-STEM students. The ANCOVA for the LA-STEM COHORT comparison showed a significant effect favoring the LA-STEM group with $F = 22.78$, $df = 1, 3269$, $p < .0001$.

The same comparison for the LA-STEM PARTICIPATION group showed a similar main effect with $F = 15.2$, $df = 2, 3269$, $p < .0001$. The pattern of post-hoc results found the students staying in LA-STEM with significantly higher scores than students in either of the other groups.

In both analyses, the propensity matching variable accounted for significant amounts of variability in the outcome variable GPA ($F = 13.5$, $p < .0001$; $F = 12.5$, $p < .0001$).

Dependent Variable: CUMULATIVE GPA

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	10.587[a]	2	5.294	13.760	.000
Intercept	4279.219	1	4279.219	11122.710	.000
MATCHING	5.197	1	5.197	13.509	.000
LASTEM COHORT	8.768	1	8.768	22.789	.000
Error	1256.907	3267	.385		
Total	35832.849	3270			
Corrected Total	1267.494	3269			

a. R Squared = .008 (Adjusted R Squared = .008)

Table 10 ANCOVA table for CUMULATIVE GPA comparison, LA-STEM COHORT Group

Dependent Variable: CUMULATIVE GPA

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	13.521[a]	3	4.507	11.738	.000
Intercept	2492.851	1	2492.851	6492.683	.000
MATCHING	4.836	1	4.836	12.595	.000
LASTEM PARTICIPATION	11.701	2	5.850	15.237	.000
Error	1253.974	3266	.384		
Total	35832.849	3270			
Corrected Total	1267.494	3269			

a. R Squared = .011 (Adjusted R Squared = .010)

Table 11 ANCOVA table for CUMULATIVE GPA comparison, LA-STEM PARTICIPATION Group

The LA-STEM cohort group earned a higher cumulative GPA than the general population, even when the propensity matching score variable was held constant.

Students staying in LA-STEM had a significantly higher GPA than students who withdrew from LA-STEM, and students from the general population.

3. *Did LA-STEM students take the same number of credits per semester on the average than the general population?*

The LA-STEM students' average number of credits was significantly *less* than the general population with F = 9.5, df 1, 3279, p = .002. The LA-STEM PARTICIPATION group comparison showed a significant difference with F = 9.57, df 2, 3279, p < .0001. The post-hoc comparison showed the WITHDREW LA-STEM group accounting for the overall difference with significantly lower scores than the LA-STEM and general population groups.

Dependent Variable:AVERAGE SEMESTER HOURS

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	34.086[a]	2	17.043	5.451	.004
Intercept	73103.365	1	73103.365	23380.477	.000
MATCHING	.011	1	.011	.004	.952
LASTEM COHORT	29.707	1	29.707	9.501	.002
Error	10246.144	3277	3.127		
Total	682667.130	3280			
Corrected Total	10280.229	3279			

a. R Squared = .003 (Adjusted R Squared = .003)

Table 12 ANCOVA table for AVERAGE SEMESTER HOURS comparison, LA-STEM COHORT Group.

Dependent Variable: AVERAGE SEMESTER HOURS

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	64.124[a]	3	21.375	6.854	.000
Intercept	41037.214	1	41037.214	13159.410	.000
MATCHING	.021	1	.021	.007	.935
LASTEM PARTICIPATION	59.746	2	29.873	9.579	.000
Error	10216.105	3276	3.118		
Total	682667.130	3280			
Corrected Total	10280.229	3279			

a. R Squared = .006 (Adjusted R Squared = .005)

Table 13 ANCOVA table for AVERAGE SEMESTER HOURS comparison, LA-STEM PARTICIPATION Group.

Students withdrawing from LA-STEM showed a significantly lower average number of credit hours taken per semester than students in the LA-STEM and general population groups.

4. *Did LA-STEM students withdraw from courses more or less than the general population?*

The LA-STEM COHORT group withdrew from courses significantly less than the general population with F = 6.22, df 1 3279, p = .013. When the LA-STEM PARTICIPATION groups were compared, a significant main effect was found with F = 5.02, df 2, 3279. The post-hoc comparison showed the LA-STEM WITHDRAW group and the general population group with higher rates of withdrawal than the LA-STEM group. The matching variable was significant with F = 6.7, p = .009.

Dependent Variable:AVERAGE NUMBER OF WITHDRAWS

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.
Corrected Model	2.955[a]	2	1.477	4.838	.008
Intercept	30.426	1	30.426	99.630	.000
MATCHING	2.062	1	2.062	6.751	.009
LASTEM COHORT	1.900	1	1.900	6.222	.013
Error	1000.776	3277	.305		
Total	1438.065	3280			
Corrected Total	1003.731	3279			

a. R Squared = .003 (Adjusted R Squared = .002)

Table 14 ANCOVA table for AVERAGE NUMBER OF WITHDRAWALS comparison, LA-STEM COHORT Group.

Dependent Variable: AVERAGE NUMBER OF WITHDRAWS

Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	4.120[a]	3	1.373	4.500	.004
Intercept	16.745	1	16.745	54.879	.000
MATCHING	1.918	1	1.918	6.287	.012
LASTEM PARTICIPATION	3.065	2	1.532	5.022	.007
Error	999.611	3276	.305		
Total	1438.065	3280			
Corrected Total	1003.731	3279			

a. R Squared = .004 (Adjusted R Squared = .003)

Table 15 ANCOVA table for AVERAGE NUMBER OF WITHDRAWALS comparison, LA-STEM PARTICIPATION Group.

The LA-STEM COHORT group withdrew from courses significantly less than students in both the general population, and students withdrawing from LA-STEM.

5. *Did LA-STEM students change more or less out of STEM majors than students in the general population?*

All LA-STEM students graduated with STEM degrees. Only one student from the WITHDREW LA-STEM group graduated with a non-STEM degree, while 26% of students in the general population switched from their intended major out of STEM.

STEM MAJOR * LA STEM PARTICIPATION GROUP

			LA STEM PARTICIPATION GROUP (0-2)			
			NON-LASTEM	WITHDREW LA STEM	LA STEM	Total
STEM MAJOR	No	Count	181	1	0	182
		Expected Count	170.5	4.5	7.0	182.0
	Yes	Count	504	17	28	549
		Expected Count	514.5	13.5	21.0	549.0
Total		Count	685	18	28	731
		Expected Count	685.0	18.0	28.0	731.0

Table 16 Cross tabulation table for STEM major comparison, LA-STEM PARTICIPATION Group.

6. What are the characteristics of the WITHDRAW LA-STEM group?

Students who withdrew from LA-STEM majors were compared with students who stayed in LA-STEM and students from the general population on all possible characteristic variables. Means are reported for all three groups.

	LA STEM PARTICIPATION GROUP (0-2)		
	NON-LASTEM	WITHDREW LA STEM	LA STEM
	Mean	Mean	Mean
AGE	21.23	21.91	20.96
GENDER (percent male)	52	55	49
Percent African-American **	4	45	33
Percent White **	85	42	50
Percent Hispanic**	2	6	5
Percent Asian	6	6	7
Percent Instate	88	89	95
Percent public school **	59	85	71
ACT comp (mean)**	27.71	27.66	28.58
ACT math (mean)**	27.24	27.76	28.87
ACT English (mean)	28.44	28.76	29.38
HS GPA **	3.86	3.79	3.88
HS Enrollment	3.34	3.29	3.03
Family Income **	127448.29	80962.87	90109.86
HS RANK	21.06	27.10	23.23

** indicates significant main effect for One-Way ANOVA at $p < .01$
Shaded groups show significant group difference in Tukey post-hoc test.

Overall, students who withdrew from the LA-STEM program were more likely to be African-American, and less likely to be white than their peers in both groups. Their HS GPA was slightly lower, and they scored lower on ACT comp and ACT math tests than those who did not withdraw.