Updated Promise Fall 2016 study

Jason Whittle

7/19/2017

Introduction

This paper discusses the effects that Fall 2016 Salt Lake Community College (SLCC) Promise (Promise) had on student retention, GPA, credits attempted and credits earned. Promise covers the cost of tuition and fees when Federal grants fall short. Eligibility depends on several factors: Utah residency, recipient of a Federal Pell Grant, full-time credit hours (12-18 hours per semester), being a degree seeking student with less than 90 attempted credit-hours, meeting with a college adviser to develop a 2-year degree plan in 'DegreeWorks', maintaining at least a 2.0 GPA, and completing 70% of attempted courses.

Students only needed to apply for enrollment and fill out FAFSA. Upon submitting the required paperwork to the Financial Aid Office at SLCC they are automatically considered for Promise. If a student qualified for Promise the Financial Aid Office notifies the student of funds and additional requirements (DegreeWorks, etc.) via their student email.

Summary of results

- Promise increased the number of credits attempted (+2.7) and earned (+2.1) in Fall 2016.
- Increased course load *did not* have a negative impact on Promise students' term GPA.
- New Promise students are forecasted to achieve the 60-earned-credits mark a full semester faster than if they had never received Promise.
- Receiving Promise had no impact on Fall to Spring retention. Promise students were already very likely to retain Fall-Spring.
- If Promise is expected to have a major impact of retention, participation in the program needs to be dramatically increased.
- Earning a higher Fall GPA, taking more semester credits and receiving Pell funds all had positive impacts on Fall to Spring retention. However, all of these are related to the requirements of Promise, which makes it difficult to attribute retention uniquely to the effect of funds distributed by Promise.

Results

Promise students completed more courses, without a decrease in academic performance[1] *and were less financially burdened.* Promise incentivizes full time "traditional" progress through college, thus could be viewed as a completion program.

Impact of Promise on credits attempted and credits earned

Both credits earned and attempted showed statistically significant increases even when controlling for year to year trends. Credits attempted because of Promise increased by 2.7 credits (+/- 0.74 at the 95% confidence interval) and credits earned increased by 2.1 credits (+/- 0.7 at the 95% confidence interval). The impacts of

[1] Due to Promise.

Promise were assessed using a difference-in-difference model (detalied in table 1). Demographic and academic controls were used along with a trend and treatment variables. Simulation shows that *a new Promise student will reach 60 earned credit hours a full semester faster than if they had never received Promise funds and only recieved Pell funds (4.5 semesters vs. 5.5) and a semester and a half faster than a non-Promise non-Pell student.*

Impact of Promise on student GPA

Promise students' GPA[2] fell from an average of 3.3 in Fall 2015 to 3.0 in Fall 2016. This result was significant with a two-sided t-test. Simple averages do not tell the full story however, since controls for a potential fall in non-Promise student GPA over the same period of time need to be considered. Progressing from introductory or remedial courses to more advanced courses could potentially reduce student GPA over time and still be a sign of progress. When applying appropriate controls there appears to be no significant negative effect of Promise on student GPA when controlling for the trend of all students enrolled in Fall 2015 and Fall 2016. The decrease in average GPA from 3.3 to 3.0 was not isolated to Promise recipients and thus is not an effect of Promise.

Where Promise didn't seem work

At best Promise had a small impact on retention but even this isn't clear. Participation in the promise program was relatively small.

Impact of Promise on Fall 2016 to Spring 2017 retention

The Fall 2016 Promise students were retained to Spring 2017 at a very high level, 83%. Pell only students retained at 76% and non-Pell qualifying students retained at 62%. However when other characteristics of student retention were controlled for such as demographics and GPA in Fall 2016 for the control group the difference in Promise to Pell retention rates cannot be said to be caused by Promise funds.

The impacts of Promise on retention were evaluated using two methods for logistic regression: Bayesian Model Averaging (BMA) and Markov Chain Monte Carlo (MCMC). Both of these methods estimated small positive impacts on retention caused by Promise. However, none of these estimates were clearly non-zero (estimated confidence intervals included 0 and negative predicted values). At this time statistical analysis cannot say that Promise had a positive impact on Fall to Spring retention however continued modeling and data will help flush this effect out.

Many controls were used in the statistical models to better isolate Promise's impact on retention. Demographics (age, gender, racial/ethnic minority) were controlled for along with total credits, transfer credits, term credit, Fall 2016 GPA, a flag for first semester students and Pell received.[3] All of the control variables were interacted with Promise to ensure that Promise's impacts were not being picked up by and mistakenly associated with these other control variables since many of the control variables are associated with Promise's requirements.

Promise as a retention program in 2016

Promise as currently designed will face hurdles in playing a significant role in SLCC retention efforts. While Promise clearly benefited those students who participated in it, there are several potential issues with it being considered a retention program. If Promise is intended to have an institution level impact on SLCC it was far too small to accomplish this (544 students in Fall 2016). The design of Promise might explain

[2]Since GPA is not normally distributed the transformation of (Term GPA - Cumulative GPA) was used for analysis.

[3]Pell award and term credit might be seen as problematic since they are critical to Promise. Several models were run that did not included them or attempted to transform them into a less problematic form. None of these transformations produced "better" retention results than what is presented in this paper.

some of the limited student participation in Promise in Fall 2016 (elaborated below). The marketing and dissemination of information about Promise may also have had an impact on the under-utilization of Promise. However, this aspect of Promise's implementation was not evaluated.

The students SLCC already retains

One of Promise's core criteria is that a student be Pell eligible; SLCC already retains over 75% of its Pell students. Promise targets students who are the most likely to retain regardless of the increased financial incentive. There are likely not many students on the fence about either returning full time or dropping out entirely who would be persuaded to stay by tuition waivers, given that these students are already receiving financial support in the form of Pell. Promise could potentially entice students to postpone transferring.[4] Since SLCC is currently the most affordable option for college in Salt Lake County it is unlikely students would transfer for financial reasons. To many Pell students, Promise might come down to a question of how many classes do they want to take.

Credit hour requirement

Money is not the only limiting factor students face in returning to college, especially for non-traditional students. If you have to take night classes for instance: how can you take more than 9 credit hours a semesters?[5] We know that "classes at the wrong time of day" is a major problem cited by SLCC students.[6] Many SLCC students might see a free $751 of classes sitting there but lack the flexibility to enroll. Promise tuition waivers are not a substitute for a full-time or even most part-time jobs. A lot of jobs don't allow you to fine tune the amount of hours you work and when those hours are. A student might correctly judge that taking 12-16 hours (to meet Promise's requirements) is too much of a commitment in addition to working, family responsibilities and other aspects of life they give priority to. Given the high number of SLCC students who have to work this might be a large road block in the path to increasing the number of students who take advantage of Promise. If Promise remains a program with participation in the hundreds per term it is highly unlikely to "move the needle" on institution wide retention rates.

Maybe not enough incentive

Promise lowers the financial burden of college but SLCC is already a cheap option and potentially not the most significant financial commitment in a students' life. The average amount of promise funds "paid" to a student was $751.70 and the maximum was $1546. That is probably not enough of a bargin to, in some students' minds, justify the extra commitment and rescheduling of life that would be required to qualify for Promise (12 credit hours) or put off other academic interest such as early transfer. Promise encourages students at SLCC to take more classes (not necessarily retain). Financially, it might be more useful to forgo Promise and just keep taking 6-9 credit hours a semester if that allows you to maintain the work hours required to pay for other necessities.

Methodology and data preparation

Data preparation

For the difference-in-difference models students who were enrolled in both Fall 2015 and Fall 2016 were used for comparisons (11,125 students). This allowed for additional controls to be placed on these models and

[4] Currently there is not enough qualitative analysis of SLCC early transfer(outs) to know the reason they leave is tuition and fees or for some other reasons.

[5] Let alone finding 12 credit hours of classes in your major at night for multiple semesters.

[6] Survey results from spring 2017.

further isolate the impacts of Promise. For the logistic regressions run to answer questions about retention the data were limited to Fall 2016 students who 1) did not receive an award in Fall 2016, 2) recorded a GPA in Fall 2016, and 3) had a record for demographic information on gender and ethnicity (21,350 students).

There were 235 *future* Promise students enrolled in SLCC in Fall 2015[7] a year prior to the start of Promise. Focusing on these students allows for comparisons of their college behavior before and after Promise for questions relating to academic performance and credits taken per semester. To separate the trend from the treatment effects for Promise students a difference-in-difference model (DiD) was used for analyzing the impact of Promise on GPA, credits attempted and earned. Various demographic and academic controls were used within the DiD model to isolate the impact of Promise. DiD analysis sheds light on how Promise students have potentially changed because of the implementation of the program. Two DiD models were estimated with both frequentist and Bayesian linear regressions.

Difference in difference model: GPA

To evaluate the impact of Promise on student Fall 2016 GPA a difference-in-difference model was used (Angrist and Pischke 2009). To evaluate the impact on Promise student's GPA, all SLCC students who were enrolled in both Fall 2015 and Fall 2016 were included. When controlling for the "trend" in GPA (what non-promise student GPAs did) the difference-in-difference model showed no significant negative impact of Promise on GPA. It appears that students enrolled in both Fall 2015 and Fall 2016 had their GPAs decline and so the 3.3 to 3.0 average decline for promise students wasn't a result of Promise. The full model results are summarized in Table 1.

Difference in difference model: Credits attempted and earned

The methodology used to evaluate the impact Promise had on student credits attempted and earned is identical to that used for GPA. A difference-in-difference model was used to isolate treatment from trend. The DiD model for both credits earned and attempted find a statistically significant increase due to promise. The estimated effect of Promise on credits attempted was a 2.724 (0.355) increase for promise students compared to other students. For credits earned there is an estimated increase of 2.111 (0.372) as a result of the Promise program. Both results are statistically significant. Both models' results are summarized in Table 1.

Retention models evaluated thus far

BMA was the primary statistical method used to model Fall to Spring retention. BMA is a robust methodology that provides robust estimated values of variables for a variety of regression based methodologies. Since retention is a yes or no variable (Did the student enroll in the next semester?), a logistic modeling methodology was used. BMA more accurately captures the uncertainty associated with statistical modeling than most other statistical methodologies. BMA evaluates all possible models given the data set[8] and then provides a weighted average estimate for the independent variables based on the probability of the sub-models. This method directly takes into consideration model uncertainty and also provides the tools for analysis of this model uncertainty. Almost all other methodologies completely ignore model uncertainty or perform specification searches that are not documented or transparent.

Robustness: Retention

Ignoring model uncertainty can dramatically under estimate the uncertainty of the statistical results generated. Specification searches introduce bias and leads to over-fitting models[9]. This may explain why BMA finds the

[7]These are the students who would go on to receive Promise in Fall 2016.

[8]2,097,152 models for this study.

[9]Including more "explanatory" variables than should be in the model. Over-fitting often leads to the illusion of a better fit and obscures "true" effects. Documented in Wang, Zhang & Bakhai 2004.

Table 1: Difference-in-difference models

	Dependent variable:		
	change in GPA	Credits attempted	Credits earned
	(1)	(2)	(3)
Trend	-0.269^{***}	-0.592^{***}	-0.757^{***}
	(0.01)	(0.05)	(0.06)
Promise students	0.049	1.786^{***}	1.930^{***}
	(0.07)	(0.25)	(0.26)
Promise effect	-0.077	2.724^{***}	2.111^{***}
	(0.09)	(0.36)	(0.37)
White	0.076^{***}	-0.470^{***}	-0.264^{***}
	(0.01)	(0.06)	(0.06)
Term age	-0.007^{***}	-0.132^{***}	-0.111^{***}
	(0.00)	(0.00)	(0.00)
Gender(male)	0.023^{*}	0.015	-0.022
	(0.01)	(0.05)	(0.06)
Credits attempted	0.029^{***}		
	(0.00)		
Cumulative GPA		0.417^{***}	2.094^{***}
		(0.03)	(0.04)
Intercept	-0.202^{***}	11.578^{***}	4.453^{***}
	(0.03)	(0.13)	(0.14)
Observations	21,704	21,704	21,704
R^2	0.043	0.113	0.189
Adjusted R^2	0.042	0.113	0.189
Residual Std. Error (df = 21696)	0.978	3.807	3.984
F Statistic (df = 7; 21696)	138.542^{***}	396.107^{***}	724.087^{***}
Note:			$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

interaction term (promise*promise) to have next to no effect and likely not a variable that should be included in the final model, while Bayesian logistic regression[10] finds it to have a small positive effect. BMA rewards parsimony in models and thus often will produce sparse models, indeed, null models are not uncommon.

Robustness refers to a parameter's ability to maintain statistical significance while other elements of the model and data are changing. The estimated value of a variable is similar across the model space: plugging in a covariate or taking one out does not effect the estimated value. BMA confronts the desire for robust parameter estimates by considering all model specification and using all of that information through weighted averaging to generate parameter estimates.

BMA results: Retention

Figure 1 displays the image plot of the BMA for retention. Red indicates a positive effect of the variable on the outcome (retention), blue indicates a negative effect on the outcome, and the beige color indicates the variable was not included. BMA considers all possible models[11], shown along the x-axis. The first model (labeled 1 on the x-axis) contained three variables indicated by the three red bars in the first column (Fall GPA, Term credits, and Pell received). These three variables were included in all the models with significant explanatory power, providing strong evidence as to their impact on Fall 2016 to Spring 2017 retention. The amount of explanatory power an individual model possesses is visually indicated by the bar length on the x-axis. Model 1 has the highest posterior probability of being the "true" model as indicated by its length on the x-axis (this is a visual representation of the weight placed on an individual variable in the averaging process).

Besides Fall GPA, Term credits and Pell award, no other variables had a strong impact on retention (as indicated by the fact they are not picked in most models.). The *Promise*Pell* (pr.pell) interaction variable in Figure 1 is showing signs of a "bad" variable as it has a different sign in different models, positive in one model and negative in two. The *promise* term has a probability of inclusion of only 13.8% (86.2% probability that *promise* has zero effect). To have weak evidence, the probability would have to be over 50%. Interacting variables (multiplying them together) allows for an analysis of variables that might have a compounding impact on each of the variables or whether one variable's relationship with the outcome variable depends on the levels of the second variable. Since Promise requires the student to be receiving Pell funds the interaction needs to be included in the model to disentangle Promise and Pell impacts. Promise will be interacted on most of the independent variables in the study.

The *Promise*Promise* (pr.pr) interaction term has a similar result with a probability of inclusion of only 18.4%. All the .pr terms represent the interaction terms, these are independent variables multiplied by promise. Again, interacting terms further isolates the effects of each variable. Interacting *promise* on *promise* will isolate the effect of the amount of promise given to students.

Bayesian logistic regression: Retention

In order to more clearly understand the probability distributions for variables with any explanatory power a second analysis was performed. The independent variables from the BMA (all the variables that had any color in Figure 1[12]) were placed into a Bayesian logistic regression as a model. This is more similar to the type of study that would be performed for this type of question others may have performed.[13] From this regression the probability distributions in Figure 2 (below) are generated. Given the way the data was transformed (odds ratios) for this analysis the value 1 on the x-axis represents no effect for the individual variable (indicated by the vertical line). Values greater than 1 represent a positive effect on retention and

[10] Just one model specification, no averaging.

[11] For averaging, only the first 150 models are used by default. For this study, BMA found 11 models with sufficient explanatory power.

[12] Variables that had any explanatory power.

[13] The type of statistical modeling that would have you ignore model uncertainty.

Figure 1: The image plot of the BMA for retention. Red indicates a positive effect of the variable on the outcome (retention), blue indicates a negative effect on the outcome, and the beige color indicates the variable was not included. BMA considers all possible models, shown along the x-axis. The first model (labeled 1 on the x-axis) contained three variables indicated by the three red bars in the first column (Fall GPA, Term credits, and Pell recived). These three variables were included in all the models with significant explanatory power, providing strong evidence as to their impact on Fall 2016 to Spring 2017 retention. The amount of explanatory power an individual model posesses is visually indicated by the bar length on the x-axis. Model 1 has the highest posterior probability of being the "true" model as indicated by its length on the x-axis (this is a visual representation of the weight placed on an individual variable in the averaging process). Source: Banner data, BMA package R (Raftery, Hoeting, Volinsky, Painter, & Yeung)

values less than 1 represent a negative effect on retention. These distributions were generated using a Markov Chain Monte Carlo simulation, where the posterior distribution is created by drawing from data 1000 times[14].

Propensity Score mathching

A thrid study was conducted to validate the BMA and Bayesian logit model results. This study used data that was pulled seperately from the data for both the BAM and Bayes logit. Propensity score matching was used to synthetically create controll group to compare the promise students too. Propensity scoring methods are meant to mimic the behavior of a controlled experiement. The 'treated' group (promise students) need to be compared to students who, but for randomness, could have been treated.

All students who were Pell eligible in the Fall 2016 semester were scored based on previous academic, demographic and course taking behavior.[15] Once these scores were generated, students were matched first by exactly grouping promise students and potential control student based on their entry cohort term. Second, the students were then paried to a minimum of one student based on their propensity score.[16] Finally, retention in Spring 2017 for the two student groups (treated by reciving promise funds vs. synthetic control) were compared to each other and Abadie-Imbens bias corrected standard errors and p-values were used to establish statistical significance of the effect.[17]

Conclusions of both methodologies: Retention

As can be seen in Figure 2, *promise* x *promise* (pr*pr, middle plot)[18] always has a clear positive effect. Out of the 1000 simulated draws none of them were at or below 1, implying a clear positive effect. The Bayesian logistic regression suggests that the more Promise funds a student received in the Fall 2016 semester the more likely that student was to enroll in Spring 2017 at SLCC.

Bayesian logistic regression reinforces the result of the BMA that *promise* has a significant portion of its distribution below 1, indicating that receiving Promise by itself in the Fall of 2016 has an unclear impact on retention (just getting Promise didn't change your likelihood of returning in the spring). *promise***pell* was a "bad" variable in the BMA and is also singled out by the Bayesian logistic regression, over 875 of the 1000 draws were at 1 implying zero effect(row 2, column 1 in Figure 2). Fall GPA, term credit and Pell award were all clearly positive with this methodology as well. Given that BMA is more parsimonious than the one model used for the results in figure 2 it seems prudent to say there is very weak to no evidence of Promise having an effect on Spring 2017 retention.

[14] This is similar to thinking about determining if a coin is 'fair' by flipping a model based on the observations of that coin 1000 times.

[15] The treatment model used a generalized boosted model for the classification problem.

[16] Matching on the propensity scores was calipered to less than or equal to .25 standard deviations. This means that propensity scores were not matched to one another unless they were within .25 standard deviation from each other. If scores failed to meet this criteria they were excluded since no acceptable match could be found (Rosenbaum and Rubin 1985).

[17] Abadie and Imbens 2002.

[18] Interaction terms explained earlier.

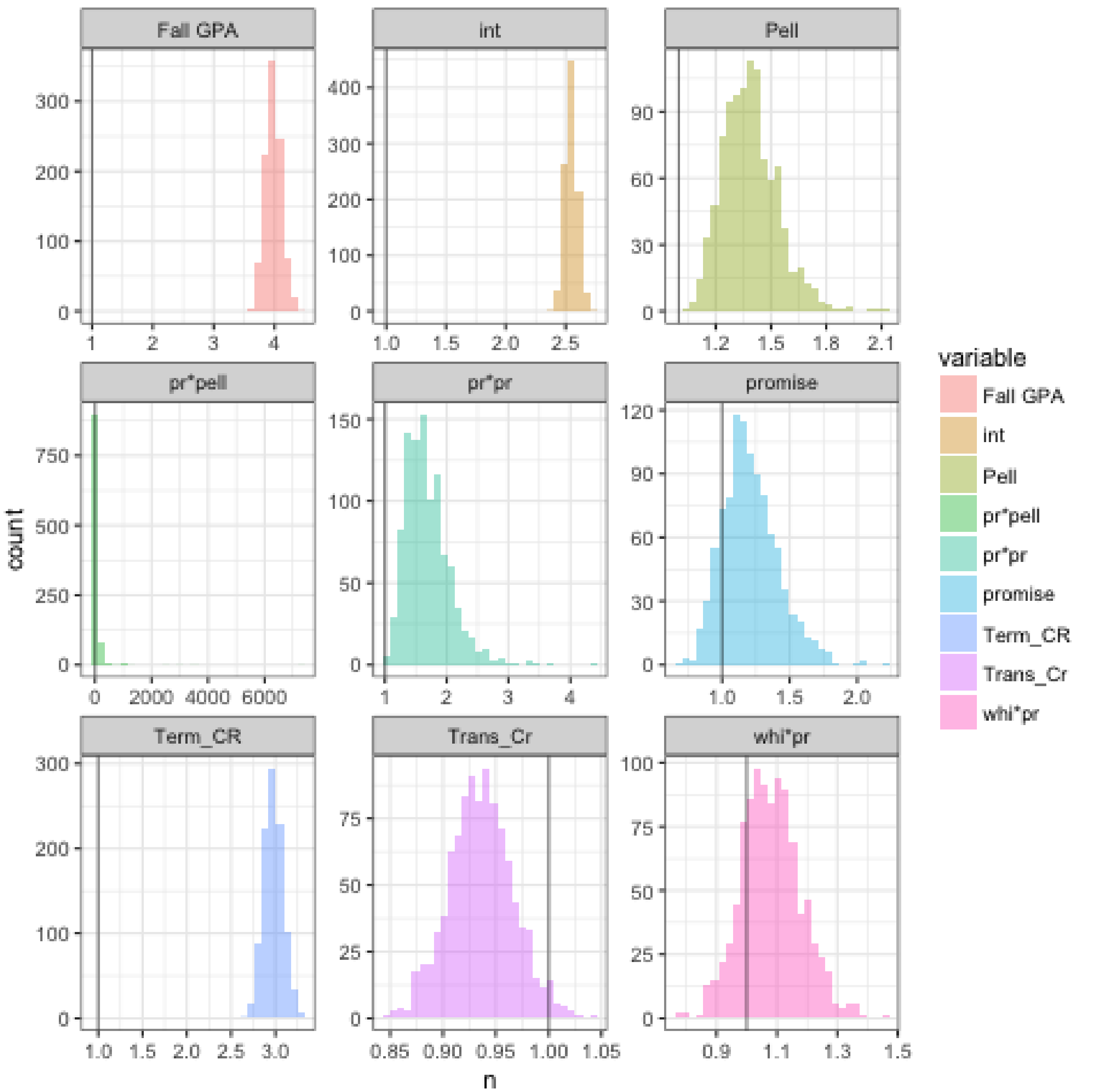

Figure 2: Posterior distributions for variables included in the Bayesian logistic regression. The vertical line indicates 1 on the x-axis or 1:1 odds (no impact). Posterior distributions that are completely to the right of the vertical line have a clear positive impact according to the model (Fall GPA, Pell, Term credits, and just barely pr*pr). All other variable posterior cross the vertical line indicating a large amount of uncertainty of the effect of these variables (it is not even clear if they have a positive or negative impact on retention).

Need for a retention study

BMA output summary

Table 2: BMA model: Retention

	p!=0	EV	SD	model 1	model 2	model 3	model 4	model 5
Intercept	100	1.31029	0.04050	1.306e+00	1.316e+00	1.298e+00	1.360e+00	1.307e+00
Age	0.0	0.00000	0.00000	.	.	.	.	.
Gender	0.0	0.00000	0.00000	.	.	.	.	.
White	0.0	0.00000	0.00000	.	.	.	.	.
Total crs	0.0	0.00000	0.00000	.	.	.	.	.
Transfer crs	21.2	-0.04041	0.08478	.	.	-1.911e-01	.	-1.881e-01
Fall GPA	100.0	1.74162	0.06845	1.743e+00	1.732e+00	1.758e+00	1.724e+00	1.747e+00
Promise	29.1	0.08491	0.22128	.	1.332e-01	.	8.217e-01	1.314e-01
Pell	100.0	0.39715	0.11289	3.527e-01	4.952e-01	3.376e-01	4.654e-01	4.779e-01
Term crs	100.0	0.76133	0.12863	8.128e-01	6.502e-01	8.284e-01	6.826e-01	6.683e-01
1st sem	0.0	0.00000	0.00000	.	.	.	.	.
age*Pr	0.0	0.00000	0.00000	.	.	.	.	.
male*Pr	0.0	0.00000	0.00000	.	.	.	.	.
white*Pr	4.8	0.01909	0.09563	.	.	.	.	.
tot*Pr	0.0	0.00000	0.00000	.	.	.	.	.
trns*Pr	0.0	0.00000	0.00000	.	.	.	.	.
GPA*Pr	0.0	0.00000	0.00000	.	.	.	.	.
Pr*Pell	10.8	0.26276	1.18867	.	.	.	4.278e+00	.
Pr.1st	0.0	0.00000	0.00000	.	.	.	.	.
GPA*1st	0.0	0.00000	0.00000	.	.	.	.	.
Term crs*1st	0.0	0.00000	0.00000	.	.	.	.	.
nVar				3	4	4	5	5
BIC				-4.976e+04	-4.976e+04	-4.976e+04	-4.976e+04	-4.976e+04
post prob				0.461	0.170	0.159	0.067	0.053

Table 2 above displays the raw BMA logistic regression output. Estimated values (EV) and standard deviations (SD) are in log-odds ratio format. p!=0 provides the probability that the variable is non-zero and should be included in the model. Most variables lack 'robustness' and therefore are not included or have a p!=0 score well below 50%.

Data sources

The retention data for this study was initially pulled by Lisa Daines from Banner and Financial Aid tables. Marie Taylor contributed additional Banner data for the GPA, credits attempted and earned studies. Data cleaning and analysis was performed by Jason Whittle.