

DRAFT: Does Use of SLCC's Student Writing Center Improve Student Performance?

Jeff Webb

*Office of Institutional Research and Reporting,
Salt Lake Community College*

March 7, 2017

Introduction

The Student Writing Center (SWC) at Salt Lake Community College (SLCC) offers peer and faculty writing tutoring for students at a number of locations: Redwood, Jordan, South City, West Valley, Library Square and the Academic Literacy Center (in TB on Redwood campus). Students using the SWC are seeking consultation on papers they are writing for a variety of courses: English 1010 and English 2010, primarily, but also courses such as Math 1030, Political Science 1100, History 1700, Communication 1010 and Psychology 1100, among others.

A previous study showed a large improvement in grades for students who used the SWC compared to students in the same classes who did not use the SWC. However, that study did not control for self-selection bias. It is possible that improved grade outcomes for students choosing to use the SWC was due to their greater motivation and engagement compared to peers. The present study updates the previous one by controlling, as much as possible, for self-selection bias. Research questions included: 1. Do students who use the SWC get higher grades in the originating course? (By "originating course" we mean the course for which the students sought SWC tutoring help.) 2. Do students who use the SWC have higher rates of retention into the subsequent semester?

Talking Points

- Propensity score matching (PSM) was used in this study to correct for self-selection bias among students who chose to use the SWC. The purpose of PSM is to make an observational study resemble a randomized study as closely as possible, thereby eliminating pre-existing differences among students as possible explanations for differences in group outcomes.
- After PSM, statistical analysis showed that:
 1. Students who used the SWC earned higher grades in the originating course compared to those who did not use the SWC. On a 4 point numeric scale using the SWC was associated with, on average, a .3 to .6 advantage in grade performance.
 2. Students who used the SWC had higher average retention rates into the next semester compared to those who did not use the SWC, by a margin of about 1-4%.

Data

Clint Gardner, Program Manager of the SWC, provided SLCC's Office of Institutional Research and Reporting (IRR) with a dataset of 26,317 unique SWC visits by 8473 students over the past 5 years (spring 2012 to fall 2016). This dataset was joined to Banner data to add student demographic and performance variables. From Banner IRR also developed a comparison group: students who did not visit the SWC but took the same courses as those taken by SWC students. In this research, we focused on the six courses for which students most often visited the SWC: ENGL 1010, ENGL 2010, HIST 1700, COMM 1010, MATH 1030 and PSY 1100. After subsetting and joining, the dataset included 3065 student course enrollments associated with a

SWC visit and 47,118 student course enrollments in the same courses without a SWC visit. Details on data preparation and cleaning are provided below.

Data analysis methods

Results from observational studies can be misleading for the reasons mentioned above. In recent years, however, methods have been developed that lessen the gap between observational studies and the gold-standard of research, the randomized clinical trial (RCT). In short, the methodological goal of an observational study should be to organize the data to resemble an RCT as closely as possible (Rosenbaum and Rubin, 1983). In an RCT, assignment to the treatment condition is unrelated to the outcome of interest because assignment is random. Not so in an observational study. But corrective options exist, the best known of which is PSM. In this study students who used the SWC were matched with students who did not, on a variety of characteristics possibly associated with use of the SWC. Matching should eliminate systematic differences between the groups, as in an RCT. Note: matching variables should pre-date the intervention, which, in this case, was SWC use.

IRR matched students who used the SWC with those taking the same courses who did not use the SWC. There was no performance information for new students (except for accuplacer scores, which are generally recognized as flawed indicators of aptitude and preparation). For this reason, the analysis was restricted to students who used the SWC in their second, or later, semester at the college. For convenience, we label students who used the SWC as the “treatment group” and those who did not use the SWC as the “control group.”

Course grade and retention to the next term were the outcome variables in the analysis. Retention was analyzed by comparing treatment and control using logistic regression. Grade differences were evaluated by comparing treatment and control using both Ordinary Least Squares (OLS) regression and ordinal logistic regression. In the former case grades were converted to a number: A = 4, B = 3 and so on. In the latter case, grades were left as alphabetic and categorical.

One problem with using OLS regression in this instance is that a number of regression assumptions are problematic: model errors, for example, are not normally distributed due to the curious “U-shaped” grade distribution at SLCC, with point masses at 0 and 4, in addition to the fact that the range of the outcome variable, grades, is not unbounded. For this particular application—comparing conditional grade means between treatment and control—OLS likely works just fine. Nevertheless, we used the additional model listed above—ordinal logistic regression—for comparison, which should be more robust to the distributional peculiarities of grades and can help confirm the results obtained by OLS regression.

Data cleaning and preparation

As noted the dataset provided by Clint contained 26,317 unique SWC visits by 8473 students over the past 5 years (spring 2012 to fall 2016). This was joined to Banner data to obtain student demographic information. Filtering choices: - First semester students were excluded for purposes of matching, as described above, and only visits to the SWC involving ENGL 1010, ENGL 2010, HIST 1700, COMM 1010, MATH 1030 and PSY 1100 were considered.

- We counted only one visit per student per term for a given class. Thus, students who visited the SWC twice for different classes were counted twice, and students who visited the SWC for the same class in two different semesters were counted twice.
- Students who did not earn a letter grade in a class (including I, W, and P) were excluded from the analysis.

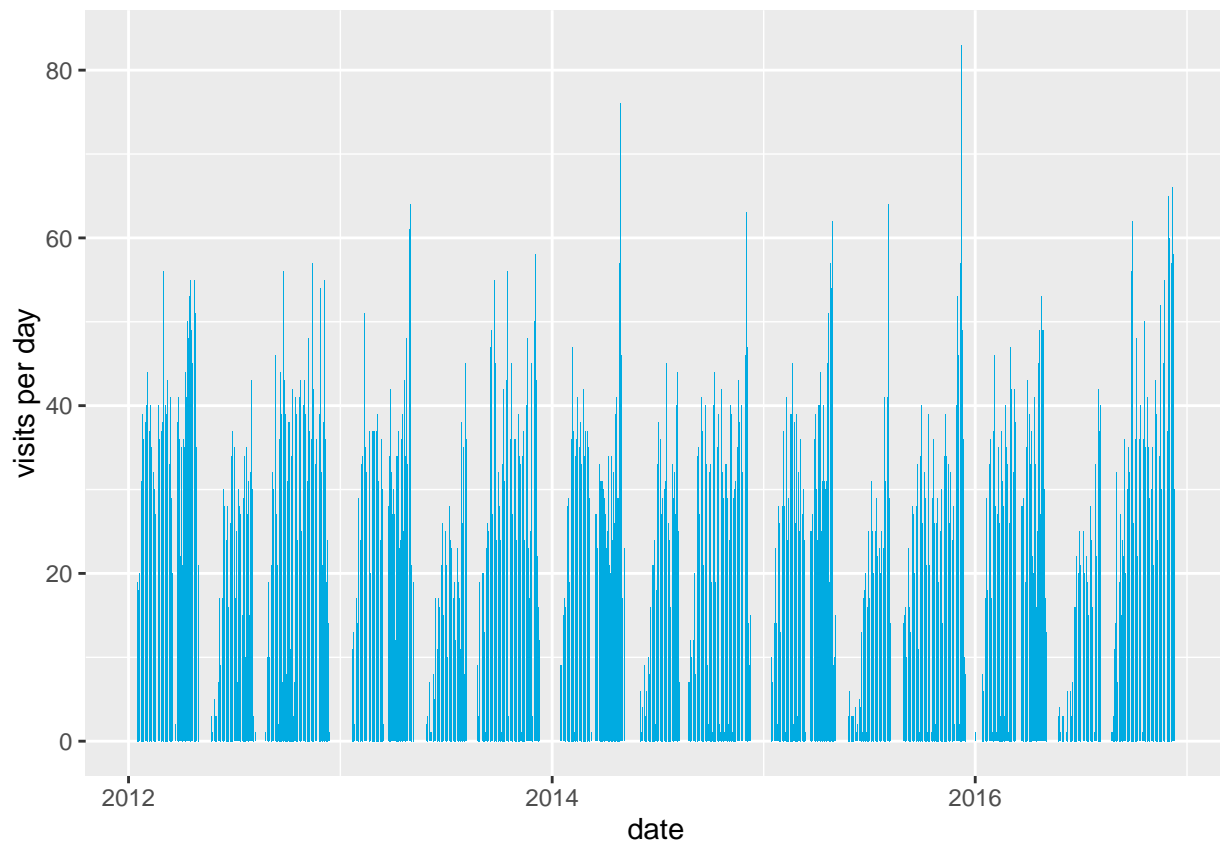
These exclusions brought the number of students in the dataset with SWC visits down to 3065 for analysis. These students were matched with those who did not have a SWC visit on the following characteristics: prior credits, prior gpa, gender, age, ethnicity (dichotomized into a binary white/non-white variable), originating course, term (e.g., 201540, 201620), and term sequence. Term sequence is a variable indicating when in a

student's career the SWC was used: second term, third term, etc. For PSM it is not necessary that matched students have exactly the same characteristics only that the *propensity* to use the SWC, defined by a logistic model using the matching variables, is *close*: matched students are "nearest neighbors."

Descriptive Results

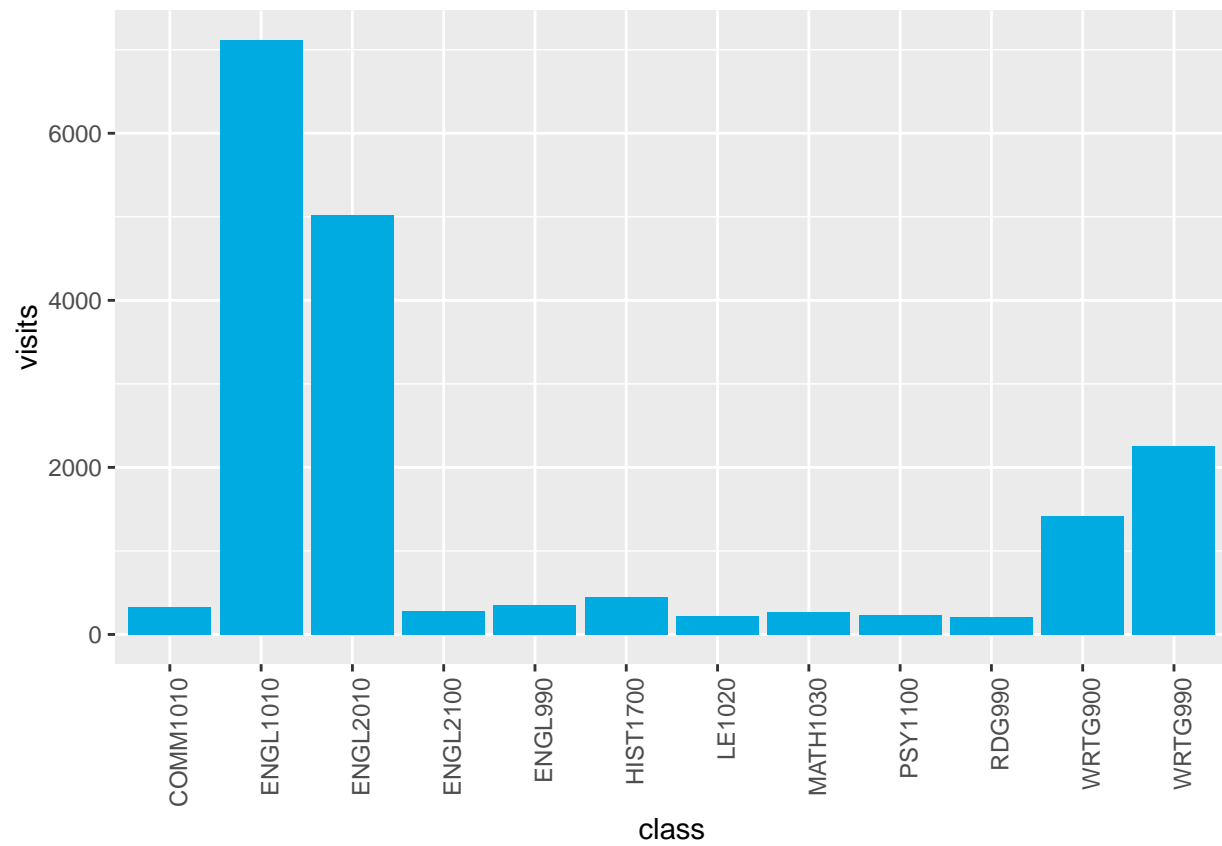
The following plots and tables present data summaries that contextualize our analysis of the relationship between, on the one hand, SWC usage and grades, on the other, SWC usage and retention. Note that, unless otherwise noted, these descriptive analyses do not use the matched datasets, and would therefore include first semester students.

Daily SWC visits by date



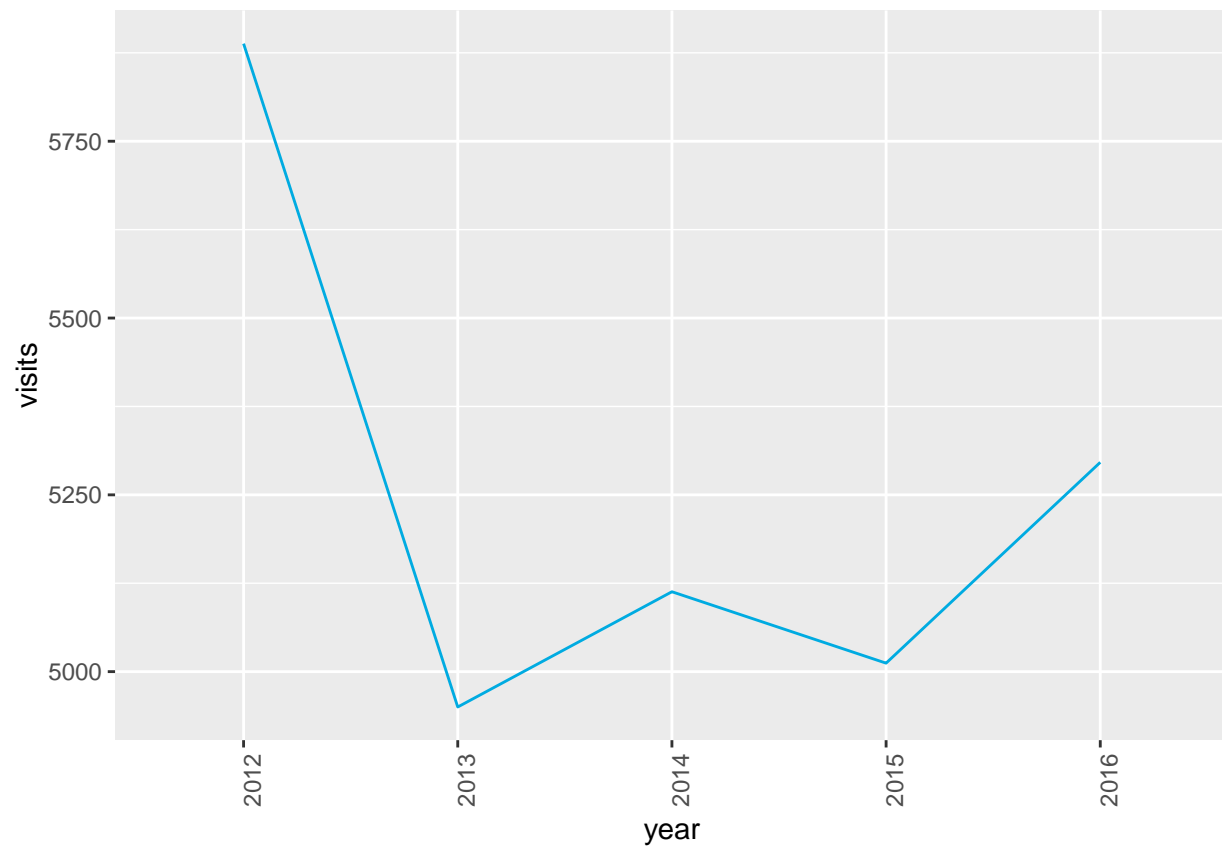
This plot makes it clear that there has been marked seasonality in the use of the SWC, with periods of closure between terms.

SWC total visits per class over 5 years



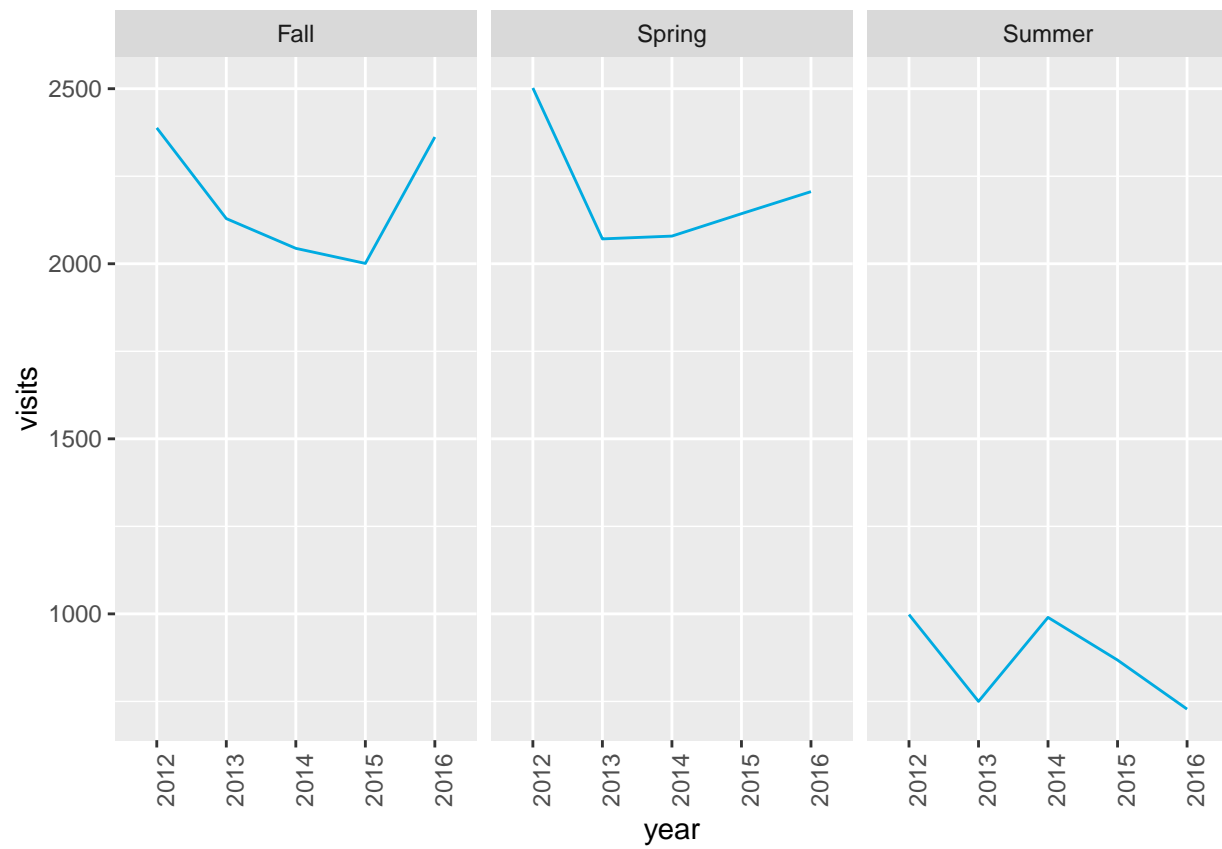
Clearly, writing classes accounted for the bulk of the SWC usage.

SWC visits by year



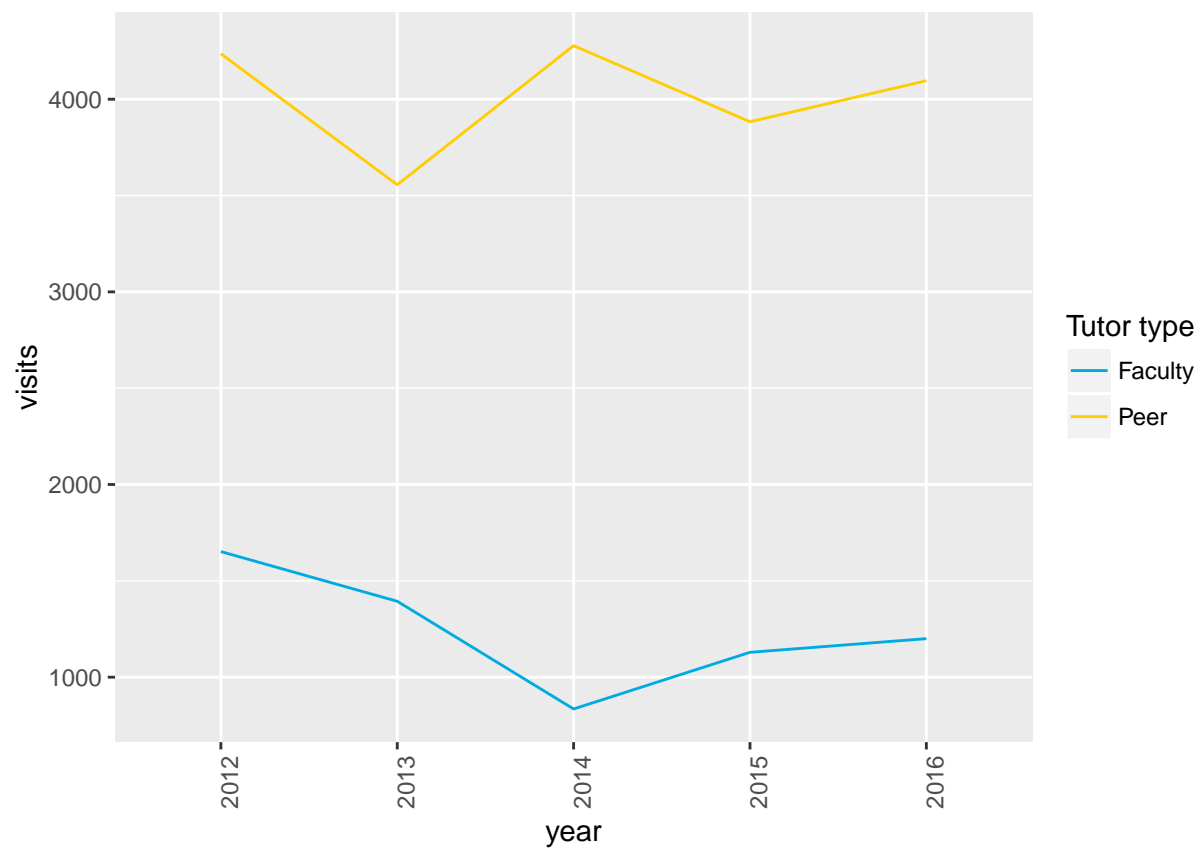
SWC visits have been slowly trending up since 2013—perhaps surprising given declining enrollments at SLCC during this period.

SWC visits by semester and year



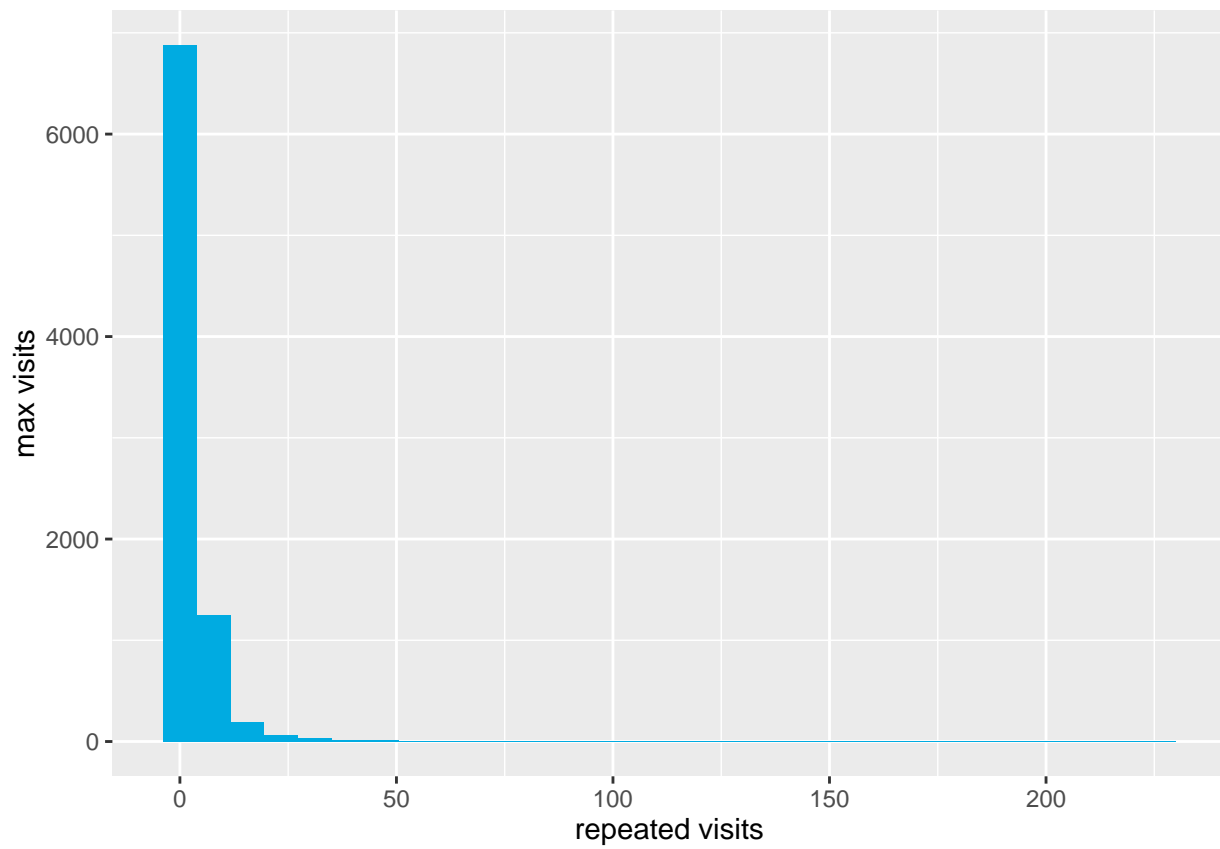
The general upward trend in SWC usage was driven largely by spring term students, until 2015, at which point usage in the fall increased dramatically also—recovering almost to 2012 levels.

SWC visits by tutor type and year



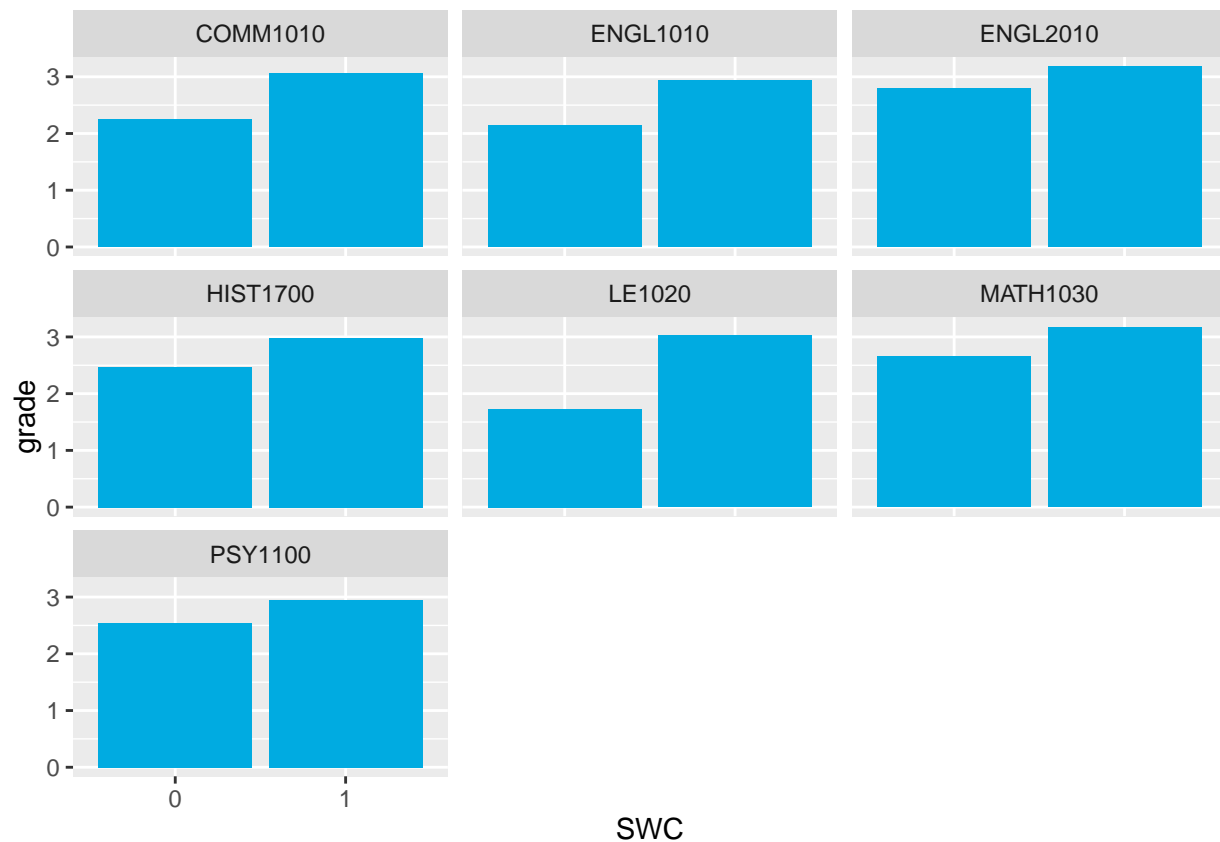
Peer tutors did the bulk of the tutoring.

Histogram of max visits per student



The vast majority of students used the SWC just once, though a small number were heavy users—100, 150, 200 visits.

Average course grades by SWC usage and class

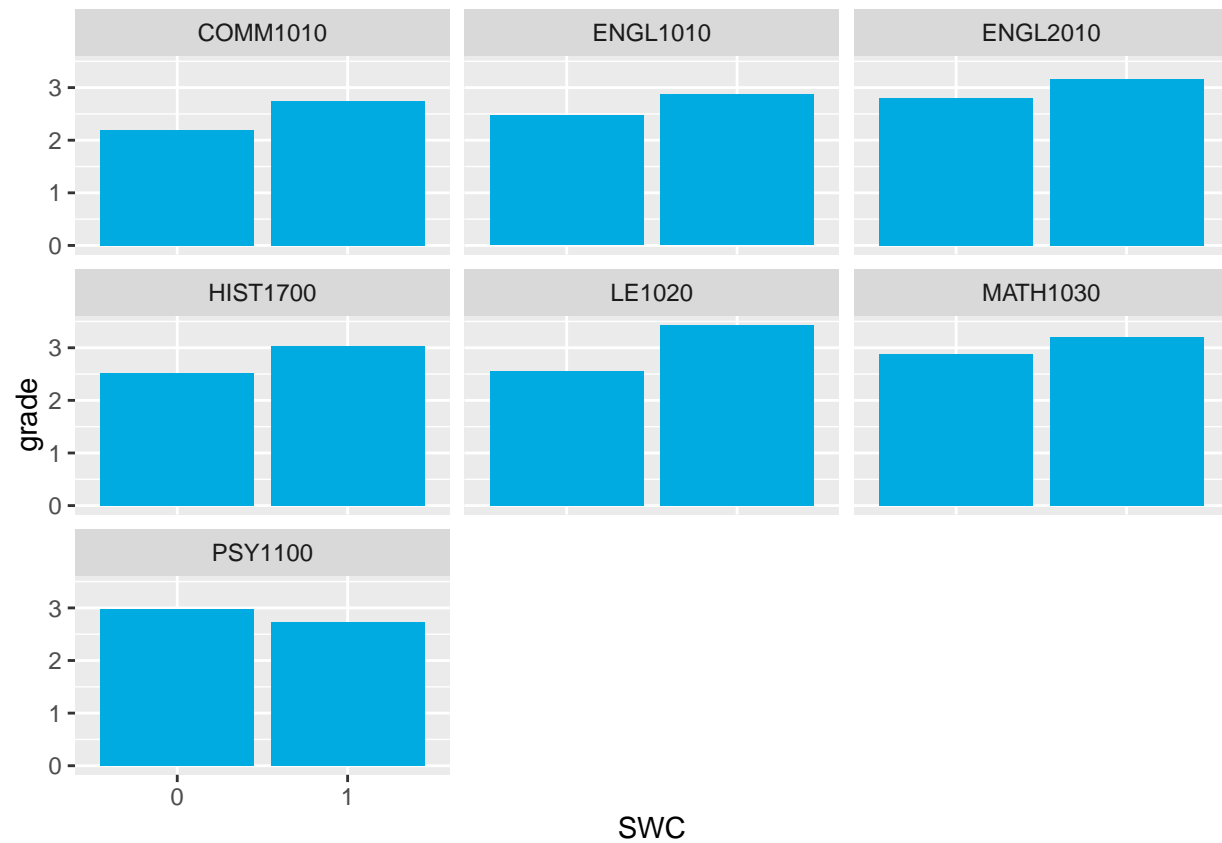


On the x-axis, 1 represents SWC usage, 0 none. Students who used the SWC got higher grades. The question addressed by this research was whether the pronounced grade effect visible in the above plot could be attributed to the SWC or whether it is explainable by pre-existing characteristics shared by students using the SWC, shown in the next figure.

Average course grades by SWC usage and class (table)

class	SWC	n	grade
COMM1010	0	5770	2.25
COMM1010	1	42	3.06
ENGL1010	0	11946	2.15
ENGL1010	1	893	2.93
ENGL2010	0	10777	2.79
ENGL2010	1	1055	3.19
HIST1700	0	8230	2.47
HIST1700	1	127	2.98
LE1020	0	1201	1.73
LE1020	1	34	3.02
MATH1030	0	1639	2.66
MATH1030	1	156	3.17
PSY1100	0	4782	2.54
PSY1100	1	47	2.95

Average course grades by SWC usage and class (matched data)

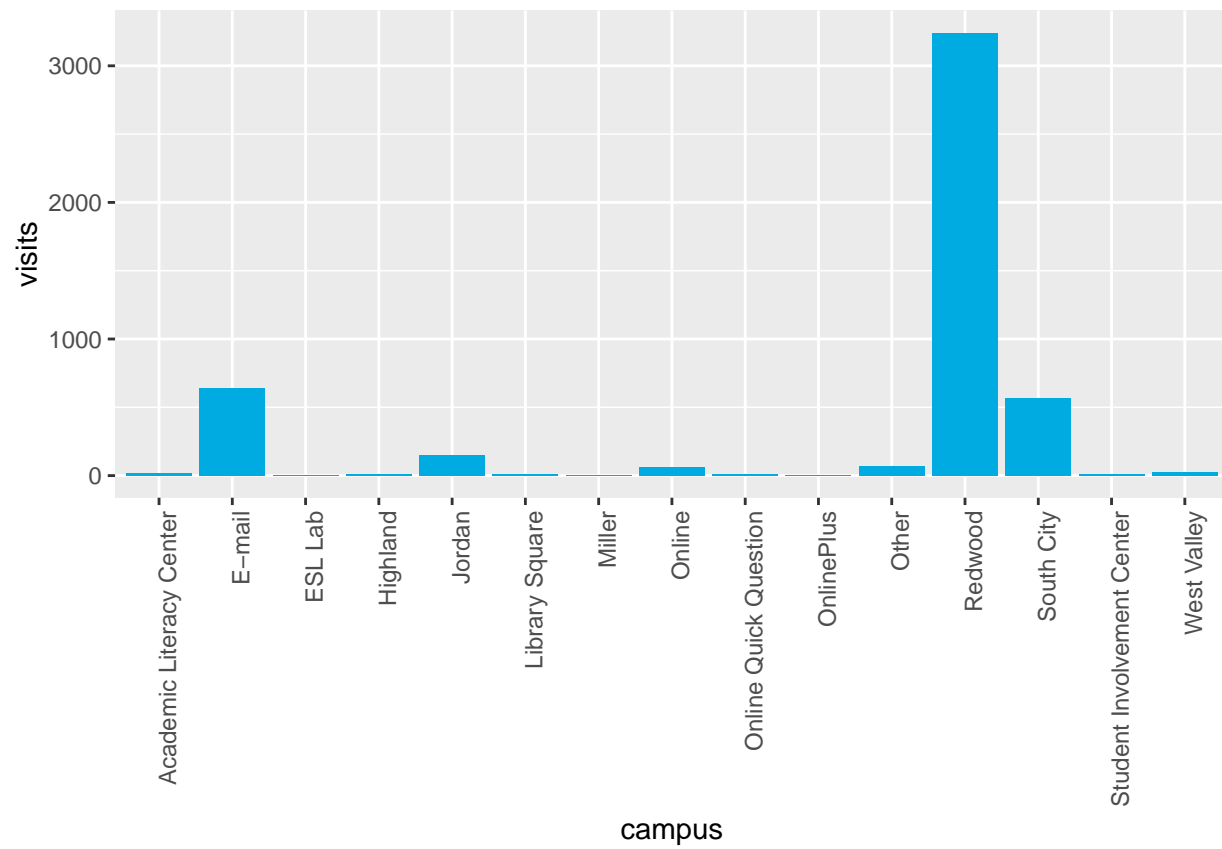


The grade effect clearly survived in the matched data. This strongly suggests that selection bias was not the source of the grade difference.

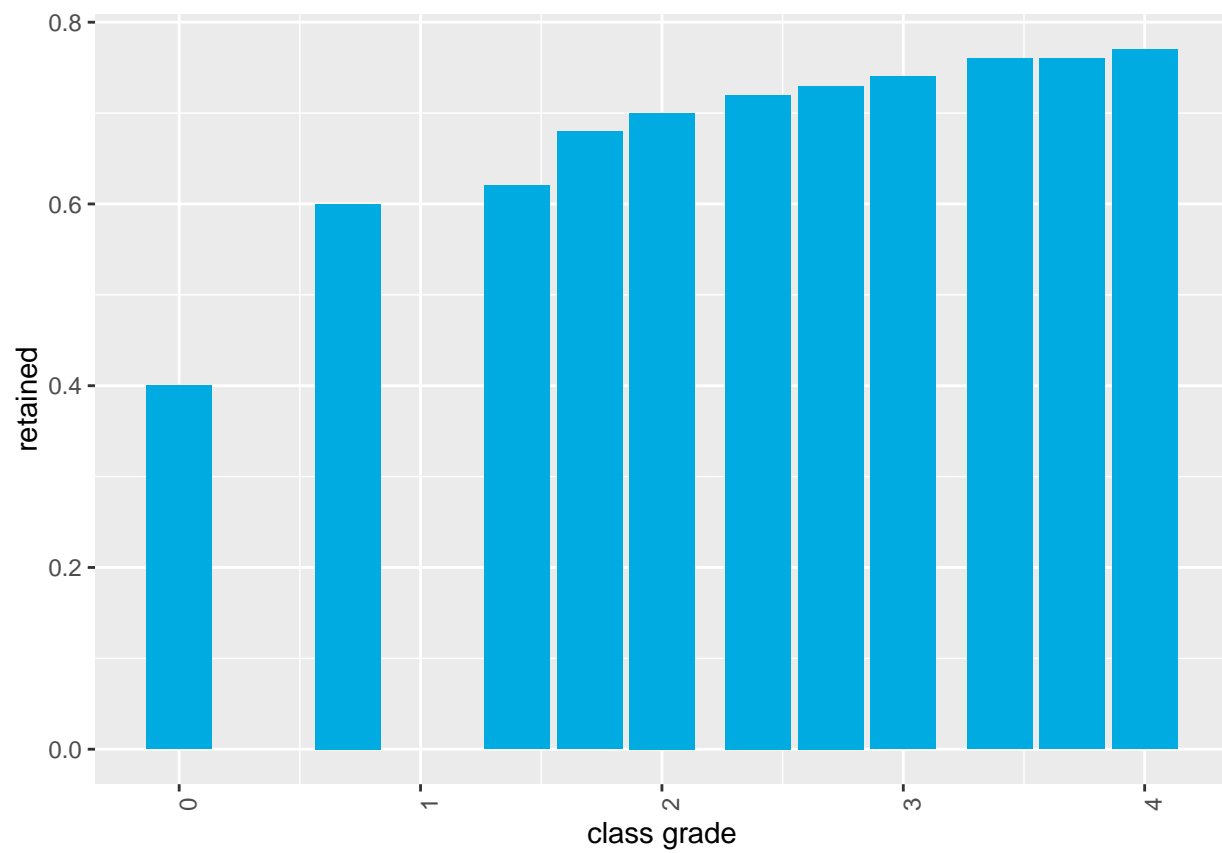
Average course grades by SWC usage and class (table with matched data)

class	SWC	n	grade
COMM1010	0	51	2.47
COMM1010	1	34	2.78
ENGL1010	0	1126	2.51
ENGL1010	1	910	2.89
ENGL2010	0	1287	2.76
ENGL2010	1	1287	3.16
HIST1700	0	114	2.65
HIST1700	1	103	3.05
LE1020	0	17	2.21
LE1020	1	9	3.31
MATH1030	0	172	2.80
MATH1030	1	154	3.21
PSY1100	0	48	3.03
PSY1100	1	37	2.75

SWC usage by campus and year

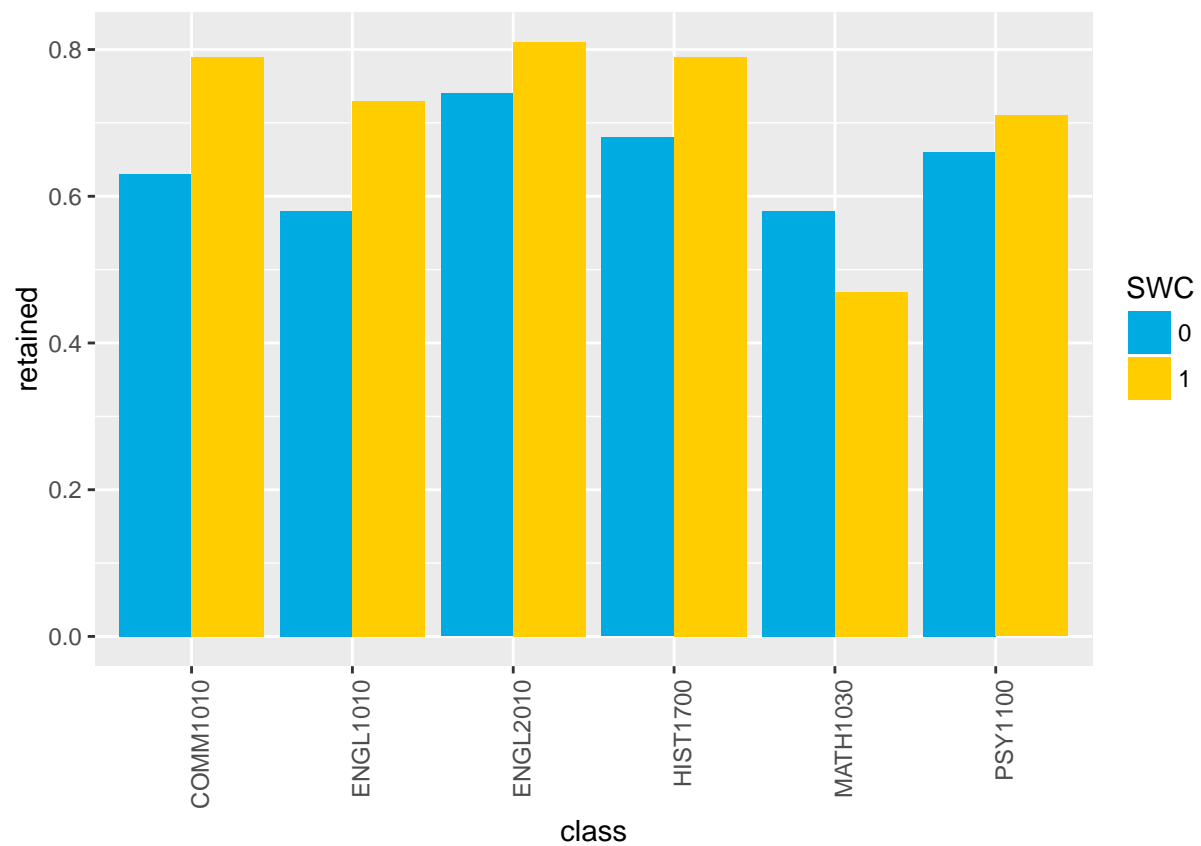


The Redwood SWC handled the most volume, followed by email consultation.

Retention by course grade

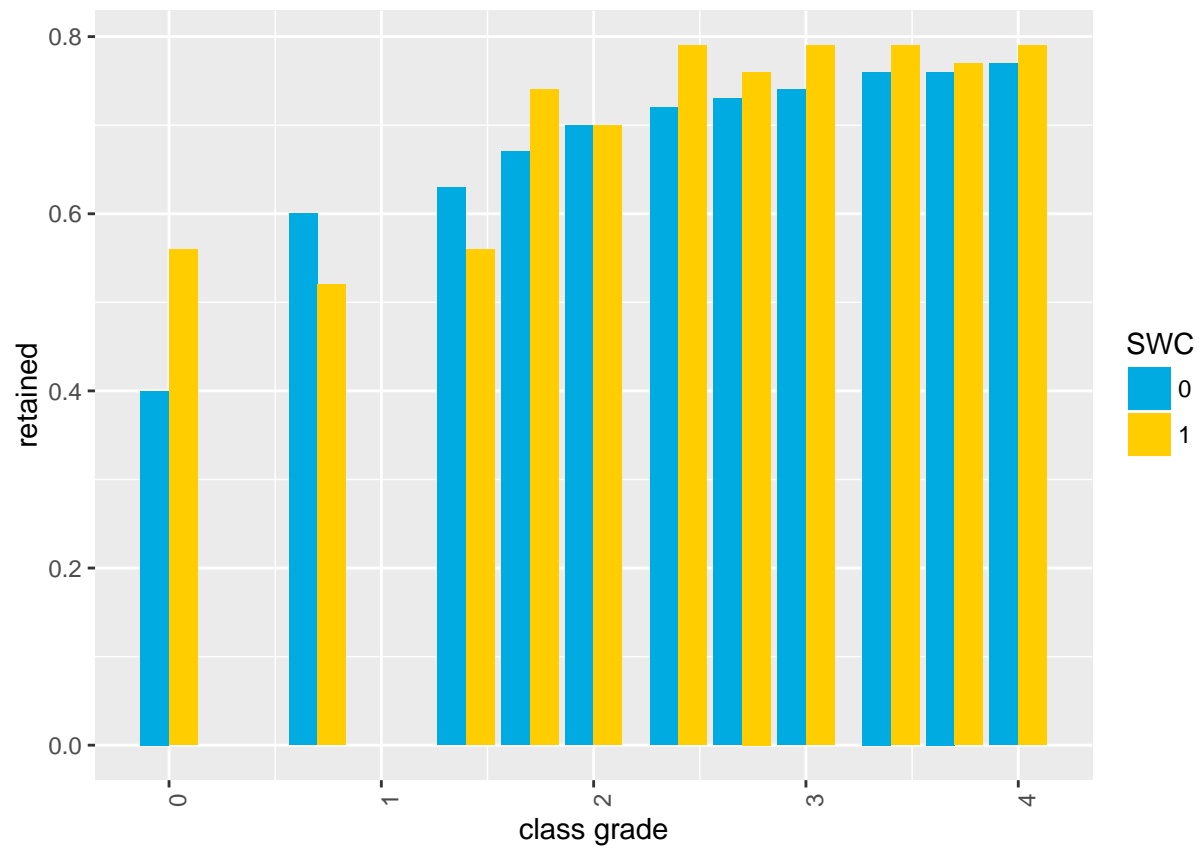
The numbering on the x-axis represents the numerical equivalent of letter grades, with A = 4 and E = 0. Retention was associated with class grade: the higher the grade, the greater the probability that a student returned the following semester.

Retention by course and SWC use



For all classes but Math 1030 SWC usage was associated with higher retention. SWC = 0 means no SWC and SWC = 1 means used SWC.

Retention by course grade and SWC use



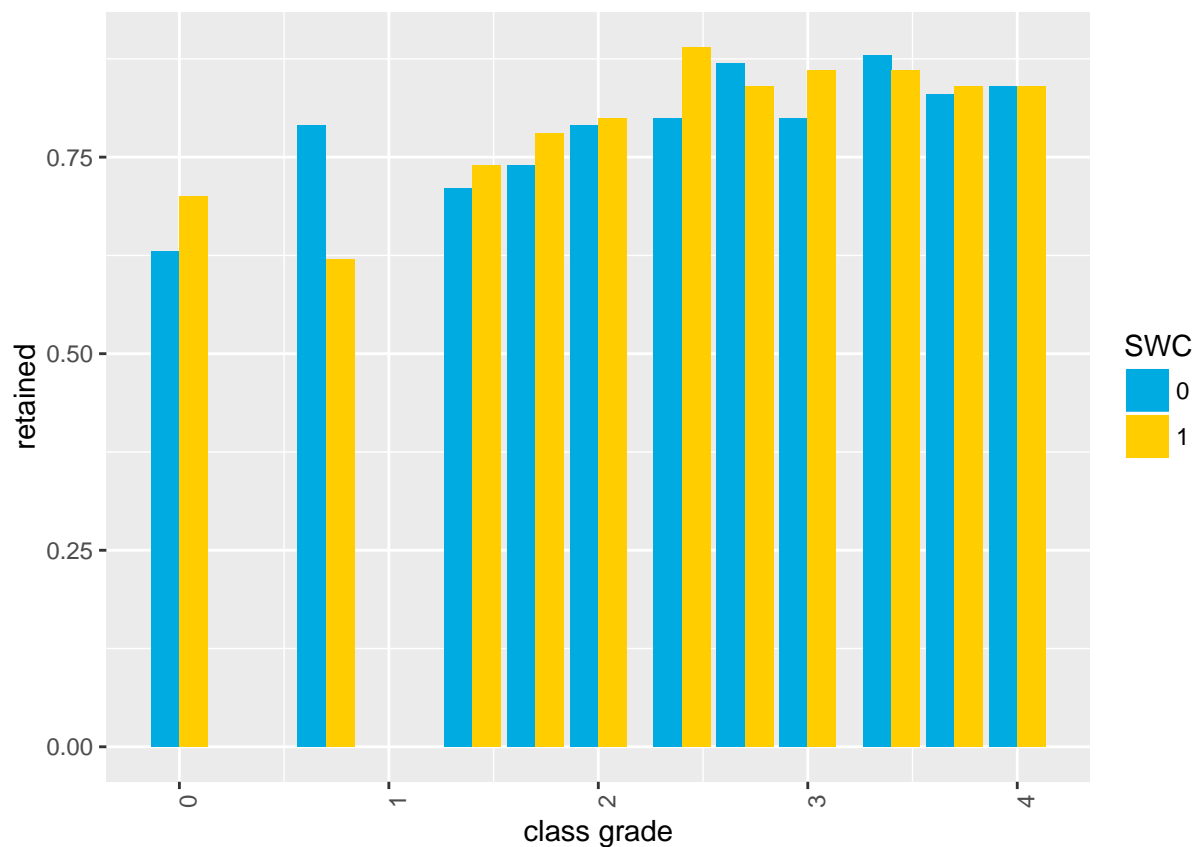
SWC was generally associated with higher retention at grades greater than C- (1.7).

Retention by course grade and SWC use (table)

class grade	SWC	n	retained
0.0	0	11179	0.40
0.0	1	245	0.56
0.7	0	615	0.60
0.7	1	21	0.52
1.4	0	521	0.63
1.4	1	27	0.56
1.7	0	1019	0.67
1.7	1	50	0.74
2.0	0	2084	0.70
2.0	1	109	0.70
2.4	0	1611	0.72
2.4	1	87	0.79
2.7	0	2359	0.73
2.7	1	115	0.76
3.0	0	4025	0.74
3.0	1	217	0.79
3.4	0	3331	0.76
3.4	1	203	0.79

class grade	SWC	n	retained
3.7	0	4728	0.76
3.7	1	340	0.77
4.0	0	11975	0.77
4.0	1	919	0.79

Retention by course grade and SWC use (matched data)



With matched data the general relationship between retention and SWC usage at higher grades was less pronounced but still largely present. The fact that the retention effect survived in the matched data, especially after controlling for class grade, suggests that selection bias was not the reason SWC students returned the following semester at higher rates.

Retention by class grade and SWC use (table with matched data)

class grade	SWC	n	retained
0.0	0	573	0.64
0.0	1	265	0.69
0.7	0	34	0.68
0.7	1	20	0.70
1.4	0	23	0.78
1.4	1	33	0.76
1.7	0	61	0.80
1.7	1	55	0.76

class	grade	SWC	n	retained
2.0	0		133	0.81
2.0	1		115	0.80
2.4	0		108	0.90
2.4	1		89	0.89
2.7	0		155	0.86
2.7	1		151	0.83
3.0	0		293	0.84
3.0	1		252	0.86
3.4	0		215	0.86
3.4	1		213	0.87
3.7	0		330	0.82
3.7	1		376	0.84
4.0	0		877	0.85
4.0	1		959	0.84

Statistical Results

Grades

Students who used the SWC received higher grades in the originating course compared to their peers who did not use the SWC, even after adjusting for differences due to gender, ethnicity, prior academic performance (gpa), accumulated credits, term, and number of terms attended. The different models—linear and ordinal regression on matched sets of students—estimated a substantial grade advantage for students who used the SWC in the matched data.

Specifically, the 95% confidence intervals for the average impact of SWC usage on grades was [0.31, 0.43] for linear regression and [0.38, 0.56] for ordinal logistic regression.

There is always uncertainty in statistical estimation. 95% confidence intervals quantify uncertainty by identifying a high-probability range for, in this case, the estimated difference in grades between the two groups, SWC and non-SWC. A range that does not include 0—meaning no difference—is equivalent to a statistically significant difference at the $p < .05$ level. In reporting two possible confidence intervals, associated with the different models, we introduce additional uncertainty, which we can summarize by saying that, on average, SWC students received higher grades compared to non-SWC students by as little as 0.31 (the lower bound of the smaller interval) or as much as 0.56 (the upper bound of the larger interval), on a four point scale, all other things being equal. This result is, we emphasize, statistically significant at the $p < .05$ level.

It is hard to know which model, linear regression or ordinal logistic regression, to trust more. While conservative estimates are generally more realistic with small samples, this is not a small sample. And given the non-normal distribution of grades, it might well be the case that the ordered logistic model fits the data better than the linear model. Consequently, the higher estimate of the grade effect may be more trustworthy. It should be emphasized that even the smaller effect size estimated by the linear model should be considered large for an educational context, where complex and numerous influences on performance tend to wash out strong relationships.

Statistical nuances aside, the main analytical result is unequivocal: Students who used the SWC got higher grades in the originating course.

Retention

Students who used the SWC also returned for the next semester at higher rates compared to their peers in the matched data who did not use the SWC, even after adjusting for differences due to gender, ethnicity,

prior academic performance (gpa), accumulated credits, term, and number of terms attended. A logistic model was used to estimate this difference in retention.

Specifically, the 95% confidence intervals for the average impact of SWC usage on the probability of retention was [0.01, 0.04]. That is, students using the SWC had retention rates that were 1% to 4% higher than students who did not use the SWC. This confidence interval did not include 0, which means that the difference in retention rates was statistically significant at the $p < .05$ level.

It was important to include course grade in the model as a covariate. Because students using SWC got higher grades in the originating course, according the above result, and because grades were strongly associated with retention, we risked misunderstanding the retention effect if we did not take course grade into consideration. Including course grade as a covariate, then, allowed us to remove it as an influence on retention and thereby to estimate the unique contribution of the SWC. This leaves the causality of the retention effect unclear, since it may well be that the higher grades associated with SWC use mediated its effect on retention. While the causal mechanism of the SWC's impact on retention is theoretically interesting, to be sure, from a practical standpoint the question of whether the impact was direct or indirect matters little. What matters is the *impact*, which is statistically indisputable.

While retention rates were associated with the grade earned in the originating course, the retention *difference* did not vary statistically with course grade. The retention difference between SWC and non-SWC students was about the same regardless of whether students earned an A, B, or C in the originating course.