Sailing into the Future:

A Program Evaluation of the Summer Academy to Inspire Learning

Laura Marshall

Daniel Scott

June 2012

Under the supervision of
Professor Bruce Blonigen

Department of Economics
University of Oregon

Abstract: Sailing into the Future: A Program Evaluation of the Summer Academy to Inspire Learning is a statistical analysis of the SAIL program began by the economics department of the University of Oregon. This program is designed to encourage pursuit of higher education among local low-income students. The data is divided into two cohorts: a nonrandomized older cohort, and a randomized younger cohort. To perform our analysis, Propensity Score Matching (PSM) is used to create a control group. Then, OLS and probit models are used to determine whether attending the SAIL camp pushes the subjects to become college ready. For the younger cohort, an OLS regression is used to determine how SAIL affects Grade Point Averages (GPA). Our research showed that SAIL attendance, while usually positive, is not significant from zero. This study is intended to provide preliminary results of the program and lay the foundation for future research.

Acknowledgments: We wish to thank Professor Bruce Blonigen for his continued advice and support with this analysis. Special thanks as well to Matt Coleman and Jennifer Robinette for their help with this project.

Table of Contents

Introduction

Although public education is available to all Americans, students are not always given equal opportunities to achieve academic success. Students from disadvantaged backgrounds, primarily those of low socioeconomic status, are less likely to enroll in and obtain post-secondary education. Reardon (2011) asserted that the current educational system perpetuates an attainment gap for students of lower socioeconomic status. Furthermore, according to Heckman (2004) there is "substantial evidence from economics, sociology and public policy studies [which] suggests that children from disadvantaged families are more likely to commit crime, have out-of-wedlock births and drop out of school". Minimizing the achievement disparity due to income has been the work of both federally funded programs, such as No Child Left Behind, and numerous localized efforts. Personal benefits clearly exist for the disadvantaged students if the education gap is minimized; moreover, investments into increasing educational attainment for disadvantaged youth have been shown to have large benefits to society as well. (Moretti 2004)

Created in 2005 to assist students with the ability to go to college, but who might not have the financial means, the Summer Academy to Inspire Learning (SAIL) is an academic intervention program that encourages higher education. Local middle and high school students from low socio-economic statuses attend a week long summer camp where they examine various academic fields taught by professors from the University of Oregon. The topics explored include, but are not limited to, physics, economics, journalism and performing arts. The creators of the camp believe that by exposing students to the university atmosphere, the attendees will become more comfortable with a college campus and see it as an attainable goal. This experience is meant to promote higher education and show the students the many opportunities available. The

SAIL program has also recently begun to offer mentoring, SAT preparation help, and assistance applying for schools and scholarships.

This study will statistically examine whether the SAIL program is achieving their goal of positively influencing college enrollment among low-income youth. If the SAIL program does indeed promote educational attainment, this should be seen as a desirable outcome, as higher education is associated with large social gains, reduced crime, and increases in wealth-generating opportunities. (Moretti 2004) These high social benefits are supported by evidence from Wolfe and Havemen (2001) who find that "the full social gains from additional schooling exceed—perhaps substantially—the 7 to 9 percent private rate found in the returns-to-schooling literature has important implications for public policy". If the intervention is not found to make any difference statistically, this result suggests that the resources could be used elsewhere to achieve their goal of improving college success rates for these students.

Data for this study was collected from transcripts and test scores obtained from the Springfield school district. The data collected for this analysis is divided into two groupings: an older and younger cohort. Since the older cohort was not randomized, matched pair sampling was utilized to create an appropriate control group for comparison to the treated subjects. The regression models for the younger cohort, which was subject to randomization, were performed using the Ordinary Least Squares (OLS) technique. The younger cohort was separated into four distinct groups: those who attended SAIL, those who were invited but did not attend, those who were selected for invitation but were not invited, and those not selected who serve as a control group. Due to the stratification of the younger cohort, we believe it is possible to determine the effect of the selection bias by regressing all possible combinations of treated and untreated subjects and comparing the results.

The results of this analysis found the effects of attending the Summer Academy to Inspire Learning to be insignificant for both the older and younger cohorts. However, this study does not specifically test the effect on college enrollment, as the data was limited by not actually having information on college entrance for graduates. Instead, this study looked at realized outcomes of academic achievement based on indicator variables that suggest a student will pursue higher education. Although no difference was found between the students who attended SAIL and those who did not, the program cannot be deemed unsuccessful until actual college entrance data is analyzed.

This program evaluation will add to the literature, as similar intervention programs do not involve direct faculty participation that introduces students to a university setting. Additionally, few program evaluations have utilized statistical analysis to determine the effect of the intervention. Thus, this analysis will lead to a better understanding of interventions across all ages, as well as speak to the efficacy of the SAIL program

Background Information

Taken from the Economics Department website (2011), the passage below describes the specific goals and design of the program:

> "UO's Summer Academy to Inspire Learning (SAIL) is a program led by volunteer UO faculty with the aim of increasing the number of low-income students enrolling and succeeding in college… We recruit 8th grade middle school students who start the SAIL camps the summer prior to entering 9th grade. These students' progress through a new camp every summer with a variety of academic topics, including; economics, psychology, biology, physics, physiology and journalism. SAIL camps are FREE and give students a hands on experience of what college life is like, including talks on getting into college and how to pay for it. The goal is for students to finish 4 years of SAIL with the belief that applying to college is the natural and normal next step in their lives, and with the tools to get admitted and to succeed… We work closely with administrators and teachers at the lowest socio-economic status (SES) middle schools in the local area to

> identify students who are bright enough to succeed in college, but are at risk of not realizing this potential due to low family income, parents without college educations, or related factors. We start with students transitioning from 8th to 9th grade, because they are old enough to start thinking about their academic future and are about to face critical decisions about college preparatory courses as they start high school."

Essentially, the SAIL program educates talented students from backgrounds that make them at risk of not attending some form of post-secondary education. It instills confidence and presents a university education as the natural step after high school. The program is focused on what to expect from the typical college experience and the many different possible fields of study available to them. In the most recent years, the camp administrators and volunteers also provide resources, such as counseling and tutoring, which may otherwise not be available to these students. By encouraging success, SAIL administrators hope they are creating a better future for every student who chooses to enter the program.

Literature Review

In the United States, the high school dropout rate is five times higher for students from lower income households than their higher income counterparts. (Zhao 2011) These disparities in educational attainment also continue into college where students from lower-class backgrounds are only half as likely to graduate from college as middle-class students. (Rumberger 2010) Inequality may stem from a variety of sources, including ability to afford tuition, college readiness, quality of past education, family involvement, as well as other factors. Successful pre-collegiate intervention programs address a multitude of these issues. (Gullat & Jan 2002)

Programs similar to SAIL, which offer social support to students, are thought to be effective for those from low socio-economic backgrounds because they encourage success and place expectations on the students. A study at an urban middle school found that poor students

improved their GPA relative to their non-intervened peers when a perceived increase in levels of social support existed. (Malecki & Demaray 2006) Additionally, these same students demonstrated fewer negative effects associated with poverty, such as low class attendance and higher rates of reprimanding, relative to their non-intervened peers. These conclusions suggest that SAIL participation will have a positive effect on a subject's grades if they perceive higher levels of social support resulting from the program.

Although social support is thought to be motivational for at-risk students, other research suggests that school policies may be more influential for student achievement. A study of the Promise Academy in the Harlem Children's Zone, known for its strict policies and long school days, shows that rigorous academic standards are effective for closing the achievement gap in mathematics among minorities. (Dobbie & Fryer 2011) This same study also found that community support systems, such as money for transportation, nutritious lunches, and a college-success office, were not influential for raising test scores. (Dobbie & Fryer 2011) The implications of this study suggest that a more intensive intervention with high academic standards is needed to have a statistically significant effect.

A program similar to SAIL is Upward Bound, a part of the federally funded TRIO programs established by the Higher Educational Act of 1965. Hosted by four-year colleges and targeted at low-income 9th and 10th graders, the program offers students a "variety of services during the year, including instruction, tutoring and, counseling" as well as an intensive 6 week summer program. (Myers et al. 2004) With the aim of encouraging economically disadvantaged students to enter college, Upward Bound has shown to be successful at increasing college enrollment for students with low initial expectations about their educational potential. (Myers et al. 2004) This program evaluation provides evidence that the longer a student receives treatment,

the more pronounced is the effect of the intervention. Due to varied participation rates for individual students, the authors found that selection bias likely causes the effect of additional participation to be overestimated. (Myers et al. 2004) The results of this program, which is similar to SAIL, but more intensive, widespread, and long-standing, suggest that our short-term investigation into SAIL will find the intervention to be ineffectual in increasing college success.

Other studies have examined the effect of an intervention program at earlier ages on high school graduation rates and college entrance. For example, Heckman (2004) found that: "...it is clear that both cognitive and non-cognitive skills are associated with lower rates of attrition from high school." Additionally, "The skills one learns in high school will determine how they perform in college and may in fact push a child to go to college. If the program pushes students to be interested in the sciences, and how to study for these classes, then it should be seen as beneficial." (Heckman 2004) Most authors agreed that if an intervention is performed after infancy, then the effects tend to only last for a few years, rather than a lifetime. (Heckman 2004) (For an intuitive understanding of Heckman's claim, see Appendix C.)

In addition to the age of treatment, gender may also play a role in the effect of intervention programs among youth. There is evidence that women respond more to academic support services and economic incentives than men, with women being more likely to utilize the resources and ultimately improve academically. (Angrist et al. 2009) Additional research also supports a gender difference when looking at the effect of tuition subsidies on college graduation, finding that the increase in degree attainment was most pronounced for non-White and Hispanic women. (Dynarski 2005)

Empirical Theory

Nonrandomized samples are exposed to selection bias due to unknown subject characteristics. One method developed for eliminating this bias is matched pair sampling. According to Rosenbaum and Rubin seminal paper (1983), matched pair sampling is, "a method of sampling from a large reservoir of potential controls to produce a control group of modest size in which the distribution of covariates is similar to the distribution in the treated group." Furthermore, a specific kind of matched pair sampling, Propensity Score Matching (PSM), "is an approach that allows researchers to match individuals in a treatment group to others who did not participate but have comparable characteristics. The innovation of PSM compared to other matching methods is that it develops a single (propensity) score that encapsulates multiple characteristics, instead of requiring a one-to-one match of each characteristic—simplifying matching by reducing dimensionality." (Peikes, Moreno, & Orzol 2008)

Peikes, Moreno, and Orzol (2008) claim that unless several conditions are met, propensity score matching can result in outcomes significantly different than Ordinary Least Squares (OLS), when randomization has occurred in a dataset. In their paper, the authors examined the effect of Social Security Disability payments on the amount of income an individual earns. (Peikes, Moreno, & Orzol, 2008) They found that: "Notably, the comparison group approach incorrectly estimated impacts on earnings. For example, the PSM approach estimated statistically significant positive impacts of between $1,000 and $1,200, whereas the experimental estimates were statistically significant and negative for the benefits counseling and waivers group (− $1,080 or − $1,161, depending on the model specification) and − $367 or − $455 and not statistically significant for the benefits counseling, waivers, and employment services group." (Peikes, Moreno, & Orzol, 2008)

The authors offered this explanation for why PSM differs from OLS results: "In the case of the State Partnership Initiative, PSM may have failed because the study sample volunteered to participate in the program based on factors that either cannot be observed or are not typically collected in administrative data." (Peikes, Moreno, & Orzol, 2008) Thus, the authors implicitly claim that PSM only yields correct results for the parameters when all relevant characteristics are known, and the data collected is noiseless. Unfortunately, it is impossible to know what characteristics are relevant before using the matching method or whether the data is noiseless.

Given that propensity score matching can only give unbiased estimates under very specific conditions, it is important to explicitly define what a propensity score is and when propensity score matching should be used. As stated earlier, matched pair sampling is necessary when no randomization exists in a study. To match two subjects, a function called a balancing score is used. A balancing score is a functional transformation of the covariates of a subject, ranging from the use of the covariates (variables) themselves to a propensity score. (Rosenbaum & Rubin 1983) Given that each variable represents a dimension, matching in terms of all the covariates could become difficult if many dimensions exist. The propensity score is considered the coarsest of all, as it reduces the number of dimensions for matching to one. (Rosenbaum & Rubin 1983) Thus, PSM can be used to ease the matching process. Due to the seriousness and difficulty of knowing the aforementioned assumptions, propensity score matching should only be used when high dimensionality is expected or no randomization occurred within the dataset.

After the obtaining the propensity score, the subjects can be matched in a variety of ways such that the probability of receiving the intervention is similar between the treated and untreated group. Among the options for matching are nearest neighbor (minimum distance), caliper (declaring a maximum distance), stratification (based on explicitly defined criterion), and kernel-

based matching (weighted average distance, where the weights are determined by the importance of characteristics). For example, if we believe that GPA is the best predictor of the SAIL intervention, then we would place a heavier weight on it than the other covariates. (Caliendo & Kopeinig 2008) Propensity score matching is utilized for the non-randomized older cohort.

Hypothesis

The graduation rate at Springfield High School is 58.8%, the lowest of Springfield high schools. (Palmer 2012) Comparatively, all but one student from the SAIL has graduated on time and 67.75% of students who have been in the SAIL program pursue a form of higher education. Empirically, this evidence suggests that the Summer Academy to Inspire Learning successfully encourages students to attended college who would not otherwise. It is possible that students who are on track to graduate are the ones being recruited into SAIL, and this analysis seeks to determine if this is indeed true.

Given evidence presented in the literature review and the specific aim of the program, the SAIL program should have a positive effect on college entrance. However, we do not have systematic data on college enrollment, making this study preliminary in nature. We will examine how the SAIL program impacts college readiness, as measured by variables that are predictive of college entrance. Furthermore, we will test whether the SAIL program impacts high school GPA. We assume that attendance of a SAIL camp should never deter an individual from entering college, nor negatively affect a student's GPA, relative to their non-intervention peers.

As stated earlier, the subsequent sections are divided into two analyses, one for the youngest cohort and a second for the older cohort. This division of the data was necessary because of the inherent differences in the data and the regression methods used for each group.

The younger cohort is comprised of the group that attended SAIL in the summer of 2011. These students are in their first year of high school during this analysis. Thus, the only college-readiness variable that will be used is their GPA from their first semester of the current year. The caveat for this cohort is the limited capacity of GPA to predict college success. However, the analysis of this group is still valuable to determine if there is any short-term effect on GPA that is attributable to attending SAIL before their freshman year. This younger cohort was randomized, allowing for the standard Ordinary Least Squares regression method to be used.

The second analysis of the older cohort consists of students from the classes of 2010, 2011, and 2012. This group was not randomized by their invitation into SAIL; all students recommended by their teachers were invited. Propensity score matching is used to create a control for this cohort. Then, a new dataset consisting of treated and untreated subjects is regressed on each of the dependent variables that indicate academic achievement. Three college readiness variables were determined for these students: GPA, AP class enrollment, and a grade of B or better in Algebra 2. Notably, actual data on college entrance is not available for this group, which limits the analysis and has implications which will be discussed further in the results section.

From these regressions we hope to gain insights into how the outcome variables are affected by SAIL camp participation. The models included other descriptive statistics which will be valuable in assessing the key determinants for academic success. Of particular interest are the variables of gender, socioeconomic status, and test scores, as we will be able to compare our results to the previous assertions in the literature. We hope that our results will help the administrators of the SAIL program better serve future students of the camp.

The Younger Cohort

Data Description

As stated in the introduction, the data collected for this project is in the form of school records from Springfield Middle School and permission slips to enter the SAIL program. The younger cohort is the group of students who were in 9th grade during the 2011 to 2012 school year. The treated subjects in this cohort attended SAIL during the summer of 2011, meaning that any effect that SAIL has on the students can be considered a short-run effect.

Below are the variables used for this model and the form that they take, continuous or binary, and the expected signs:

GPA: The Grade Point Average of the student their first semester in high school, based on a 0 to 4 scale. The students' GPA is the dependent variable for this model and will be interpreted also as a predictor of college enrollment, with higher GPAs suggesting a student is more likely to attend college.

Attended Sail: A binary variable that is 1 if they were invited by teachers and attended SAIL, and zero otherwise. The focus of this analysis is on this variable, which is anticipated to have a positive coefficient.

Rejected Sail Invitation: A binary variable that is 1 if they were invited but chose not to attend SAIL, and zero otherwise. A negative coefficient is expected, as these students were not motivated to accept the extra help from the program.

Recommended Only: A binary variable that is 1 if they were recommended by their teacher to attend SAIL but were not actually invited to attend, and zero otherwise. A positive coefficient is expected because the teachers recommended students who are above average academically and have the ability to go to college.

OAKS Reading: A continuous variable representing the score the student achieved on their Oregon Assessment of Knowledge reading test in 8th grade. A positive coefficient is anticipated, as reading is a primary medium of education.

OAKS Math: A continuous variable representing the score the student achieved on their Oregon Assessment of Knowledge math test in 8th grade. A positive coefficient is predicted, as those who perform well on skill tests should be more talented academically.

Gender: A binary variable that is 1 if the student is female and 0 if they are male. The literature review suggests that the coefficient will be positive, denoting that females obtain higher grades, all else equal.

Algebra or Less: A binary variable that is 1 if the student was in Algebra or a lower class in the fall term of their freshman year in high school, and 0 otherwise. Algebra is the most common freshman class at Springfield High School, also has one of the lowest passing rates. We expect a slightly negative coefficient for this variable.

Geometry: A binary variable that is 1 if the student was in Geometry, a level higher than Algebra, in the fall term of their freshman year in high school. This variable is thought to have a positive coefficient; a higher level of math correlates positively with a higher GPA.

Algebra 2: A binary variable that is 1 if the student was in Algebra 2, the highest mathematics class taken in the sample. This coefficient should be positive, and greater in magnitude than that for Geometry.

Free Lunch: A binary variable that is 1 if the student is eligible for the free lunch subsidy from the federal government. Qualification for free or reduced lunch is determined by socioeconomic status, and will therefore be a proxy for income level. A negative coefficient is

anticipated, in accordance with the literature review, which suggests there is an attainment gap for low SES students. (Zhao 2011; Reardon 2011)

Reduced Lunch: A binary variable that is 1 if the student is eligible for reduced lunch from the school system, based on family income. This coefficient is also thought to be negative, but less than *Free Lunch*, as the income threshold for entering reduced lunch is higher.

Race White: A binary variable that is 1 if the student identified as being of the race white, and zero otherwise. This coefficient could be positive or negative, but we think it is most likely positive since they are not a minority. Additionally, it is not in our reference category.

Race Hispanic: A binary variable that is 1 if the student identified as being of the Hispanic race, and zero otherwise. Hispanics make up a large share of the student population, but are not part of the reference category. We assume this variable will be slightly negative due to potential language barriers.

Attendance: A continuous variable representing the percentage of the previous year that they attended school. We expect this covariate to have a positive coefficient, as the more a student attends class, the better prepared academically he or she should be.

Minor Incidents: A continuous variable of the number of minor incidents that are on their school record. The coefficient of this variable is anticipated to be negative, as those who are being reprimanded for their behavior are not focusing on academics.

Major Incidents: A continuous variable of the number of major incidents that are on their school record. This variable is also expected to have a negative coefficient, as being forced out of the classroom should harm ones academic performance.

The data for this project was limited, as we could only obtain information from the students at Springfield High School and not those SAIL attendees who went on to other high schools. This restraint resulted in the loss of data points, leaving us with only 22 students who attended SAIL, 21 who were invited but did not attend, and 12 who were not invited but were recommended. Additionally, there were 248 students who served as a control group who were never recommended by their teacher. Following is a brief summary of the descriptive statistics for the 303 student sample:

Table 1: Data Description: Younger Cohort

	Attended SAIL		Recommended SAIL		Rejected SAIL		Control Group			
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Min.**	**Max.**
GPA	2.95	0.913	2.74	1.02	2.895	0.941	2.46	1.21	0	4
Special Ed.	0.136	0.351	0.127	0.336	0.167	0.389	0.190	0.393	0	1
Total Minor Issues	5.73	8.25	6.71	10.1	9.75	14.7	4.05	6.67	0	56
Total Major Issues	1.95	2.90	2.67	4.26	3.83	6.06	3.59	6.55	0	42
Female	0.682	0.477	0.582	0.498	0.583	0.515	0.460	0.499	0	1
Native Amer.	0	0	0.036	0.189	0	0	0.016	0.126	0	1
Asian	0	0	0.018	0.135	0	0	0.020	0.141	0	1
Black	0.045	0.213	0.036	0.189	0	0	0.008	0.090	0	1
Hispanic	0.273	0.456	0.327	0.286	0.583	0.515	0.453	0.457	0	1
Multiple	0.456	0.213	0.073	0.262	0.0833	0.289	0.077	0.267	0	1
Pacific Islander	0	0	0	0	0	0	0.008	0.090	0	1
White	0.636	0.492	0.509	0.505	0.333	0.492	0.585	0.494	0	1
Free Lunch	0.682	0.477	0.655	0.480	0.583	0.515	0.625	0.485	0	1
Paid Lunch	0.318	0.477	0.236	0.429	0.333	0.492	0.294	0.457	0	1
Reduced Lunch	0	0	0.109	0.315	0.083	0.289	0.081	0.273	0	1
Previous Attendance	94.6	4.47	94.6	5.01	95.6	4.44	93.2	6.855	56.9	100
OAKS Math	235.8	7.65	235.3	6.71	234.8	6.56	234.8	9.50	209	272
OAKS Reading	235.1	7.71	233.5	6.75	231.9	6.88	233.5	8.25	215	254

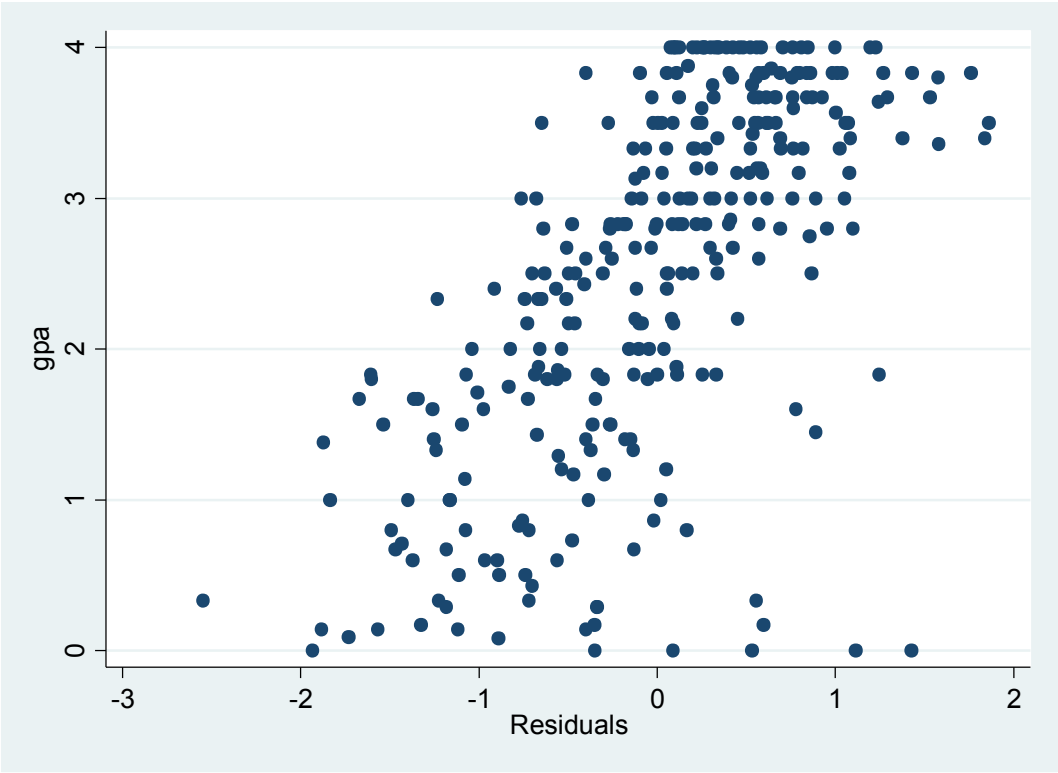
The descriptive statistics exhibit that there is a large Hispanic and white population at Springfield High School. A majority of subjects are eligible for free or reduced lunch program, implying that the sample is representative of the population. The differences between subjects recommended by their teachers and their peers are minimal. However, there are approximately 10% more females in the recommended group. Additionally, the GPA of the control group is approximately 0.3 units less (-11.4%) than those recommended and 0.5 units less (-19.9%) than those who attended SAIL. Overall, the OAKS scores are similar across groups, suggesting that any GPA differences are not the result of differences in knowledge or testing ability. The gender composition of the groups is significantly different. There is a slight minority female in the control group (48.6%) but a large majority (68.2%) for the students who attended SAIL. It is important to note that there are not any students on reduced lunch for the group which attended SAIL.

There are limitations to this data set. For example, it is unclear if GPA is a quality predictor of college success because the ultimate goal of the SAIL program is enrollment in higher education, not improved high school grades. Geiser & Santelices (2004) found that high-school grade point average is "consistently the best predictor" of freshman grades in college and four-year college outcomes as well, suggesting that using GPA as a predictor is appropriate for this analysis. However, due to only one term of high school being completed, GPA is an imperfect approximation of its future self. Additionally, other unobservable differences may exist which are unaccounted for, creating the possibility of omitted variable bias.

Methodology

Regression analysis is utilized to determine the effect that attending SAIL had on the grade point average among the 9th grade sample. The cohort was randomized by who was invited to attend the camp, with the controls for this group include those not invited, but recommended, by the teacher; those who rejected the invitation; and those who were never recommended. Due to randomization, the use of Ordinary Least Squares as a method of regression is appropriate. A linear model was used and thought to be the best specification of the model.

Using a model with the previously described variables, the Breusch-Pagan / Cook-Weisberg test for heteroskedasticity found the Prob> Chi² to be 0.0317, suggesting the presence of heteroskedasticity. Graphically, it appears as:



To correct for heteroskedasticity, the model was transformed by dividing each variable by σ, assuming heteroskedasticity of the form σ^2. To obtain sigma, the squared residuals from

the estimated model were then regressed on the independent variables. Then, the fitted values were estimated from that regression and sigma is the absolute value of the square root of the fitted values. This corrected model was the final model used for the regression of the younger cohort.

Results

The results of the primary model took into account all attainable variables that are thought to be exogenous to GPA. That is, for the ith student, we estimate:

$$GPA_i = \beta_1 + \beta_2 SAIL_i + \beta_3 RejectSAIL_i + \beta_4 RecommendedOnly_i + \beta_6 X_i + u_i$$

where X_i is a list of control variables and u_i denotes the disturbance term. Table 2 gives the coefficient estimates with varying inclusion of control variables and different combinations of SAIL experiences. The coefficient estimates represent the estimated effect of that covariate on GPA, holding all other variables constant. Standard deviations are in the brackets below the coefficients; asterisks represent the respective levels of statistical significance.

Table 2: Effect of SAIL on GPA for Younger Cohort

	GPA			
Variables	**(1)**	**(2)**	**(3)**	**(4)**
Attended Sail	0.0674251	0.0442886	0.0171416	0.0986467
	(0.1644324)	(0.1642391)	(0.1725512)	(0.1692666)
Rejected Sail Invitation	-0.0628917		-0.0798082	-0.0222686
	(0.2085031)		(0.2208139)	(0.2144363)
Recommended Only	0.4968726		0.3010647	0.5834939
	(0. 2841728)*		(0.2919655)	(0.2920579)**
Oaks Reading	-0.0094894	-0.0112284	-0.0107438	0.01846
	(0.0075154)	(0.007503)	(0.0076934)	(.0006772)***
Oaks Math	0.0178818	0.0179623	0.0300181	
	(.0076935)*	(0.0077328)**	(0.0078362)***	
Gender	0.3800753	0.4018959	0.6281483	0.251693
	(0.1235278)***	(0.1237398)***	(0.1217788)***	(0.1213458)**

Algebra or Less	-0.3879175		-0.4738836	-0.4071951
	(0.118657)***		(0.1234808)***	(0.1206352)***
Geometry	-0.3726715	0.0155252	-0.4891416	-0.2201543
	(0.1725065)**	(0.2062127)	(0.1817254)***	(0.1632849)
Algebra 2		0.3938081		
		(0.1214703)***		
Free Lunch	-0.2103785	-0.2277439	-0.2198821	-0.2206955
	(0.1068771)**	(0.107103)**	(0.1135725)	(0.1100554)**
Reduced Lunch	-0.0857964	-0.1040667	-0.0372117	-0.1158816
	(0.1820356)	(0.1812389)	(0.1931485)	(0.1874201)
Race White	-0.2440055	-0.2430472	-0.2709548	-0.1974559
	(0.1230383)**	(0.1232422)**	(0.1305144)**	(0.1236724)
Race Hispanic	-0.2431183	-0.2208302	-0.1716114	-0.1737857
	(0.1460531)*	(0.1453595)	(0.1546804)	(0.1491746)
Attendance	0.0269522	0.0273213		
	(0.008503)***	(0.0085208)***		
Minor Incidents	-0.0217037	-0.0187807		-0.0174267
	(0.0088084)**	(0.00858)**		(0.0090044)*
Major Incidents	-0.0284329	-0.0296451		-0.0416041
	(0.0116512)**	(0.0116612)**		(0.0116106)***
R^2	0.8952	0.8934	0.8801	0.8896
Adjusted R^2	0.8898	0.8886	0.8751	0.8847
Root MSE	1.1734	1.1797	1.249	1.2003
Prob > F	0	0	0	0
Observations	303	303	303	303

***** = Significant at the 1% level**
**** = Significant at the 5% level**
*** = Significant at the 10% level**

Note: The control groups for these regressions were students who were never recommended by their teachers to attend SAIL, males, and those who were not eligible for free or reduced lunch. Additional controls varied by regression.

The variable of interest, *Attended SAIL*, was not found to be statistically significant in any of the regressions performed. The regression results suggest that the SAIL camp does not produce short run effects that are reflected in attendees' GPAs. It may be the case that a student must attend multiple years of the camp for a measurable effect.

Also of interest are the coefficients for the *Rejected SAIL Invitation* and *Recommended Only* variables. Rejecting a SAIL invitation did not have an apparent effect on the student's grade point average, contrary to our predictions that these students might be less motivated than those who attended. Although it was not statistically significant, the coefficient for the rejected

variable was negative in all variations of the regression. The variable representing students which were recommended but not invited, *Recommended Only*, did have a positive coefficient in regression (1) that was also significant at the 10% level, and in regression (4) which was significant at the 5% level. This finding suggests that those who were invited have a higher GPA than the rest of the sample of 9th grade students, which was expected based upon the goal of recruiting able students into SAIL.

The independent variables in the model had similar effects to those of other studies. Our data supported the proposed gender difference in education that suggests males achieve lower grades and are more often in remedial classes than females. (Angrist et al. 2009) In accordance, the results of this analysis find that there is a positive effect on GPAs if the student is female, statistically significant at the 1% level in all regressions. The proxy for socioeconomic status, the *Free Lunch* variable, also supported evidence that low income students perform worse in school relative to their peers. (Zhao 2011) This variable was estimated to have a negative effect on GPA, significant at the 5% level in regressions (1) and (2). Math also seemed to be a predictive of higher GPAs, both OAKS math scores and being in Algebra 2 have statistically significant positive coefficients. In regression (4) *OAKS Math* and *Attendance* were removed from the regression, *OAKS Reading* became statistically significant, likely due multicollinearity, reflected in the high pairwise correlation coefficients of 0.9983 and 0.9935, respectively.

The results of this data on the 9th grade sample of students suggest that the SAIL program does not have a short-run effect on the academic success for students who attend the program. Based upon these results we recommend a further study at later ages to determine if there is a difference in actual college enrollment and graduation.

The Older Cohort

Data Description

The data collected for this project is in the form of school records from Springfield Middle School and permission slips to enter the SAIL program. The older cohort is the group of students who are either graduated or currently in the 12th grade during the 2011 to 2012 school year. The treated subjects attended various numbers of camps over their high school career, with a minimum of one and a maximum of four.

Below are the variables used for this model and the form that they take, continuous or binary:

Algebra 2: A dependent variable, it is a binary variable that is 1 if the student received a grade of B or higher on average in Algebra 2, and 0 otherwise. This variable is interpreted as a proxy for college readiness, as those with higher mathematical ability should be more college ready than their peers in lower levels of math.

GPA: A dependent variable representing grade point average. This variable is bounded between zero and four. GPA will be interpreted also as a proxy for college enrollment based on achievement, with higher GPAs suggesting a student is more likely to attend college.

AP Course: A dependent variable, it is binary with 1 corresponding to a student having taken at least one Advanced Placement class and 0 if they have not been enrolled in any. This variable is a measure of college readiness, as those who pursue AP courses are thought to be revealing their preference for college enrollment.

SAIL: A binary variable that is 1 if they attended the SAIL camp, and zero otherwise. The focus of our analysis is on this variable, which is anticipated to have a positive coefficient.

OAKS Reading: A continuous variable representing the score the student achieved on their Oregon Assessment of Knowledge reading test in 8th grade. The anticipated coefficient for this variable is positive, as greater reading comprehension should promote college readiness.

OAKS Math: A continuous variable representing the score the student achieved on their Oregon Assessment of Knowledge math test in 8th grade. This coefficient is expected to have a positive coefficient, as those with greater mathematical ability should be more college ready compared to their peers.

Gender: A binary variable that is 1 if the student is male and 0 if they are female. The review of the literature suggests that this will have a negative coefficient because females are more inclined to succeed academically.

Free Lunch: A binary variable that is 1 if the student is eligible for free lunch from the school system. This classification is determined by the Federal Government, and is based on preset income levels. The coefficient for this variable is expected to be negative.

Reduced Lunch: A binary variable that is 1 if the student is eligible for reduced lunch from the school system. This label is determined by the Federal Government, and is a higher income range than free lunch eligible students. This variable is expected to have a negative coefficient.

Paid Lunch: A binary variable that is 1 if the subject pays for his or her lunch and 0 otherwise. This variable is expected to have a positive coefficient, as it is indicative of higher income levels.

Race White: A binary variable that is 1 if the student identified as being of the race white, and 0 otherwise. This variable is expected to have a positive coefficient, as whites tend to enroll in college at higher rates than all other races. (US Census 2008)

Race Hispanic: A binary variable that is 1 if the student identified as being of the Hispanic race, and zero otherwise. This variable is also to have a slightly negative coefficient due to potential language barriers.

Race Multiple: A binary variable that is 1 if the student identifies as being Multiple races, and zero otherwise. This variable is in our reference category, and is expected to negatively influence the constant term.

Attendance: A continuous variable representing the percentage of attendance for their high school career. This variable is expected to have a positive coefficient, as those who attend class should have a higher GPA relative to their non-attending peers.

The next table is a comparison of the differences in means, which also describes the data used in our analysis.

Table 3: Differences in the Mean Among Matched Subjects

	Treated Subjects		**Control Subjects**		
Variable	# Observations	Mean	# Observations	Mean	Difference in means
AP course?	32	0.3125	61	0.3606557	-0.0481557
Attendance %	32	0.9661626	61	0.9607572	0.0054054
GPA	32	2.94125	61	2.86541	0.07584
HS graduation	32	0.5	61	0.4918033	0.0081967
Free lunch	32	0.375	61	0.442623	-0.067623
Paid lunch	32	0.46875	61	0.5081967	-0.0394467
Reduced lunch	32	0.15625	61	0.0491803	0.1070697*
Race-Hispanic	32	0.1875	61	0.2295082	-0.0420082
Race-Multiple	32	0.09375	61	0.0655738	0.0281762
Race-White	32	0.71875	61	0.704918	0.013832
Male	32	0.375	61	0.5081967	-0.1331967
ESL	32	0	61	0.0655738	-0.0655738
Algebra 2	32	0.4375	61	0.3442623	0.0932377
SAIL	32	1	61	0	1

*=P<0.10, **=P<0.05, ***=P<0.01

The differences in means between control subjects and treated subjects are minimal at best, except for the number of males, white ethnicity, and reduced lunch subjects. Over seventy percent of our matched pair dataset is composed of white students, which is a smaller proportion

than population as a whole. Additionally, those subjects of multiple races are underrepresented in our matched pair dataset. Furthermore, our matched sample has a lower high school graduation rate than the SHS student population. (Palmer 2012) As a note, at the time of this regression, the class of 2012 had not graduated yet, biasing this statistic downward. Therefore, we suspect that an updated dataset will reveal that the matched sample has higher graduation rates relative to their peers. The variable for reduced lunch was the only statistically significant covariate after matching for a two sided t-test. Since only eight subjects in the dataset were on the reduced lunch program, the statistical significance may be a function of small sample biases, rather than poor matching on said covariate.

Although there is not data on college outcomes for the control group, this information is available for the students in the older cohort that attended SAIL. Below is a summary of their post-grad choices. The original data detailing exact school choices is available in Appendix E.

Table 4: College Outcomes for Older Cohort SAIL Groups

	4-year College	Community College	Military	No College/Unknown
2012	19.23%	53.85%	11.54%	15.38%
2011	35.0%	30.0%	10.0%	25.0%
2010	25.0%	43.75%	0.0%	31.3%
Totals	**25.81%**	**41.94%**	**8.20%**	**17.74%**

From this data we see that the majority of the SAIL students do go on to pursue post-secondary education. However, the majority of these students are choosing to go to community college, which has lower retention rates than four-year schools and does not allow students the same opportunity of upward mobility as four-year universities. (ACT News Release 2012) Future analysis of the SAIL program will have to consider the difference between these two outcomes and what is desired by the SAIL program.

Methodology

To perform the analysis, Propensity Score Matching was used to create a match sample, which was then regressed on multiple dependent variables. As earlier stated, a propensity score is the probability of a subject receiving a treatment. Various regression techniques can be used to create a propensity score. Since a linear regression model can generate probabilities of less than zero or greater than 1, we decided to use a probit regression model, or a model developed for regressing binary (0 or 1) dependent variables. Per literature recommendations, a Stata package (Pscore) was downloaded which creates propensity scores when regressed. The package takes the probit model used, regresses the independent variables on the dependent variables, then, using the data given, inputs all information into the model. For example, if my model is of the form $I = .05 + .2White + .09GPA$, where I is the probability of receiving an intervention, White is a binary variable for race, and GPA refers to the grade point average of an individual. Thus, if a subject is white and has a 4.0 GPA, then the probability of receiving the intervention is $I = .05 + .2(1) + .09(4.0) = 0.61$.

Stata performs this calculation for all subjects in the dataset. To then match a treated and untreated subject, a technique called nearest neighbor matching with replacement was used, where nearest neighbor matching is the act of matching those who received the intervention to those who did not based on proximity. Replacement refers to allowing a single control subject to be used several times in the sample, if said subject is matched to a different treated subject each time. For an example of nearest neighbor matching, consider the following: if a control and treated subject had a propensity score of 0.50, then according to the model, the two subjects would be a perfect match. For our sample, we decided on a caliper, or maximum distance apart, of 0.01. Using a caliper ensures the control and treated subjects remain similar across control

covariates. After matching the treated subjects to their controls, a new data set is created which contains solely these students. This new data set is then used to regress independent variables, including the variable for SAIL attendance, on different dependent variables (Algebra 2, GPA, AP). From this, we can discern the effect of SAIL on these variables.

Bootstrapping was implemented due to a small sample size after matching treated and control subjects. Bootstrapping is a technique used to improve the accuracy of the standard errors in regression analysis. This is accomplished by altering the distribution of a sample to make it better match the population distribution. The results of each regression are reported in addition to the robust standard errors for comparison.

Results

The data table listed below captures the controlled effects of the intervention on the entire dataset. This model is used to create the propensity score for matching untreated and treated subjects in the older cohort.

Table 4: Results of Propensity Score Development

Dependent variable	
SAIL Attendance	
Independent Variables	**Coefficients**
Special Education	-0.264605
	(0.4986752)
Male	-0.1636381
	(0.1789526)
Free Lunch	0.4040807
	(0.1930709)**
Reduced Lunch	0.6419502
	(0.3489718)*
8th Grade OAKS	0.0369565
	(0.0111563)***
Race-Hispanic	0.331684
	(0.3304865)
Race-White	0.4032974
	(0.2849757)
Constant	-10.80996
	(2.687434)***
Number of observation:	685
LR chi2(7):	22.62
Probability>chi2	0.002
Pseudo R2	0.0875

*=P<.10, **=P<.05, ***=P<.01

From the regression, the matching groups are statistically insignificant based on race and gender. Our proxy variables for income, free and reduced lunch, are significant, which implies that low income students are likely to receive the intervention. Additionally, the OAKS test scores for reading are also significant, which may be because those who received the intervention were better at reading comprehension than those who did not. Furthermore, the constant is significant, implying that a part of our reference category, multiple races, female, and paid lunch, is more likely to receive the intervention. This could be attributed to race, as all non-Hispanic and white races constitute 15% of the dataset. Notably, our model does not include the

state testing scores for mathematics. This occurred because the variable failed to satisfy the balancing property necessary for developing a propensity score.

The results used for primary analysis are the marginal effects. These can be characterized as a change in the dependent variable with respect to a change in an independent variable. For the OLS regressions, these are the coefficients in the bootstrapped data table, which can be found below the marginal effects of the probit regression. Additionally, all other data tables are listed in the Appendix section D with a brief explanation.

Table 5A: Marginal Effects of SAIL on College Readiness using Boostrapped S.E.

Dependent Variable	Algebra 2			AP Courses?		
	(1)	(2)	(3)	(1)	(2)	(3)
Independent Variable						
Free Lunch	-0.1118821	-0.1221188	-0.1234519	-0.1634504	-0.1596621	-0.1747018
	(0.088)	(0.14245)	(0.14634)	(0.15646)	(0.10897)	(0.13186)
Reduced Lunch	-0.3037348	-0.2902621	-0.2349906	-0.2276078	-0.1147653	-0.0245174
	(0.11101)***	(0.10356)***	(0.12113)*	(0.10757)**	(0.08462)	(0.14694)
Race-Hispanic		0.2688255	0.2708478		0.9996253	0.999205
		(1.23752)	(1.27432)		(0.0007)***	(0.00092)***
Race-White	-0.0154406	0.1672898	0.1826705	0.0651632	0.9824042	0.966862
	(0.11997)	(1.12533)	(1.07481)	(0.14631)	(0.01685)***	(0.02782)***
Male	-0.1703059	-0.1735189	-0.1816693	-0.1262908	-0.1383574	-0.1443909
	(0.10033)*	(0.13677)	(0.15345)	(0.09794)	(0.09432)*	(0.10907)
ESL	0.2165613	0.23206	0.19611	0.6157282	0.9023707	0.8992229
	(0.38007)	(0.86153)	(0.87811)	(0.13092)***	(0.07043)***	(0.03333)***
SAIL Attendance	0.136812	0.1489688	0.1316083	-0.0110254	-0.004757	-0.0233477
	(0.10557)	(0.14164)	(0.14465)	(0.12261)	(0.0983)	(0.13101)
OAKS-Reading	.0240906	.0262427	.0256647	0.029984	.0285431	0.0304984
	(0.01547)	(0.01505)*	(0.01481)*	(0.01426)**	(.01368)**	(0.0175)*
OAKS-Math	-.0023829	-.0025643	-.0025953	0.0003275	0.0006048	.0006831
	(0.00604)	(0.0065)	(0.00812)	(0.00659)	(0.00674)	(0.0058)
Attendance			2.667046			3.103238
			(1.327)**			(1.60012)*

*=P<0.10, **=P<0.05, ***=P<0.01

Table 5B: Effect of SAIL Intervention on College Readiness Using Bootstrapped Standard Errors

Dependent Variable	**GPA**		
	(1)	**(2)**	**(3)**
Independent Variable			
ESL	0.5946374	0.6605178	0.5837874
	(0.3259857)*	(0.5462986)	(0.52601)
SAIL-Attended	0.1053381	0.1788603	0.1585715
	(0.1564311)	(0.1981314)	(0.1914109)
8th Grade OAKS-Reading	0.0233064	0.0281814	0.0225118
	(0.0136607)*	(0.0139819)**	(0.0127644)*
8th Grade OAKS-Math	0.0055791	0.0063791	0.0061438
	(0.0072151)	(0.009218)	(0.0091613)
Male	-0.0624697	-0.0757635	-0.0697049
	(0.1395591)	(0.1708836)	(0.2051415)
Race-Hispanic	0.8222737	0.8470163	
	(0.2895039)***	(0.3990877)**	
Race-White	0.9907826	0.8999999	0.2956551
	(0.2727768)***	(0.3026281)***	(0.1874242)
Lunch-Free	0.2161759	0.1973549	0.244588
	(0.1419336)	(0.1701535)	(0.1849302)
Lunch-Reduced	-0.1243744	-0.3473413	-0.437359
	(0.4330997)	(0.30347)	(0.3255661)
Attendance%	8.516875		
	(2.258777)***		
Constant	-13.16021	-6.245443	-4.236234
	(3.116768)***	(3.027572)**	(2.818078)
Number of Observations	93	93	93
Replications	49	47	47
Wald chi2	44.81	27.6	26.97
Prob>chi2	0	0.0011	0.0007
R-squared	0.399	0.2062	0.1484
Adjusted R-squared	0.3257	0.1201	0.0673
Root MSE	0.6651	0.7597	0.7822

*=P<.10,**=P<0.05,***=P<0.01

Across all regressions, the 8th grade *OAKS reading* variable is significant at the ninety-five or ninety-nine percent significance level. This would imply that reading ability influences AP classes, GPA, and the grade received in Algebra 2. Reading ability positively affecting GPA and AP courses is unsurprising, as a majority of students who had an AP course listed on their transcripts were enrolled in AP Literature. Undoubtedly, most classes in high school require the ability of reading comprehension, suggesting that those who are better at reading will have a higher GPA. The reading scores significance with respect to Algebra 2 is somewhat unexpected.

One potential explanation is that those who performed well on the reading portion of the OAKS test are more likely to be future college students, and need a decent grade in Algebra 2 to enter a university. Alternatively, the course could be taught with an emphasis on reading a textbook versus visual learning.

Both binary variables for race are significant with respect to GPA and AP courses. Unsurprisingly, our study concurred with the literature on race and academic achievement, as white students can still expect to earn a higher GPA on average than other races, such as Hispanics.

The SAIL attendance variable is insignificant in every model. This is not surprising, as only thirty-two subjects received the intervention, weakening the power of statistical tests. Additionally, it may be the case that the SAIL program does not improve any predictors of college. Rather, the intervention only promotes college entrance. The coefficients for attendance are positive with respect to *GPA*, *Algebra 2*, and *AP courses*. It is assumed that the amount a person attends class is endogenously related to all dependent variables. Therefore, the significance of this variable may be tied to the intimate relationship between attending class and a student's GPA, Algebra 2 grade, and AP courses.

As stated earlier, marginal effects capture a change in a dependent variable with respect to an independent variable. In the probit models, the only significant independent variable for both Algebra 2 and AP courses is reduced lunch. Not included in the income proxy variables is paid lunch, suggesting that a change from paid lunch to reduced lunch will negatively impact grades in both Algebra 2 and AP courses. That is, those who are deemed impoverished by the Federal Government are less likely to succeed academically, evidence that is supported by the literature review. (Reardon 2011) Furthermore, those who learned English as a second language

were less likely to enter AP courses than their peers. Given that AP courses are taught in English, and a majority of them are reading based, it would imply that if one is not fluent in the English language, then language is a deterrent to becoming college ready.

Conclusion

The SAIL program is designed to help those who, academically, may be capable of attending post-secondary education, but may have some impediment to reaching their goal, such as family income, motivation, or confidence. A majority of students entering SAIL are classified as impoverished by the Federal Government, as determined by free lunch eligibility, but many of them have the ability to go to college, which was desired by the SAIL administrators. Of the nineteen students who graduated in 2011, twelve chose to attended, with eight actually going to the University of Oregon. This anecdotal evidence suggests that there may be an actual effect on the decision of these students who were in SAIL to go to college. It is possible that SAIL has an effect at the margin when student actually make the decision. Unfortunately, this evaluation did not have data on actual college enrollment, which we hypothesis to be effected by SAIL. While our results are, in general, insignificant, it does not imply that the SAIL program is not meeting its goals. The data set used for the older cohort included ninety-three observations, which renders little power for the tests. Additionally, due to time and data constraints, no analysis was performed of actual college entrance. The variables used as predictors for college entrance *(AP Classes, GPA, Algebra 2*) are imperfect predictors, as not everyone with a high GPA or who takes AP courses goes on to college. Additionally, Algebra 2 may not be the best measure of college readiness, because it relates only to the specific skill of math and not other types of academic learning. Of course, even if one can manage a B or better in Algebra 2, they still may

not pursue higher education. To improve upon this ambiguity for the future, actually college entrance data should be used.

Due to the flaws in this study, we recommended additional investigation into the SAIL program. By this time, full randomization should occur within the program, eliminating the need for Propensity Score Matching. Furthermore, more detailed instructions should be given to the teachers who select students to enter the SAIL program, as our results suggest they are choosing disadvantaged students, but not all of the students selected seem to have the ability and drive to attend college. One suggestion for the SAIL program is to have the students complete a survey about their expectations about college and try to recruit the students who are doing well academically but may not be fully confident that college is possible for them. This suggestion is based off of our results that academic ability matters most and the research of Malecki and Demaray (2006) which stresses the importance of a student's personal expectations. Notably, the SAIL students also consist of some children from Hamlin Middle School, subjects we could not use in this evaluation because of data limitations. To minimize this issue for the future, we recommend only having students from one school for each SAIL camp. This alteration would also allow the SAIL program to invite students from other schools in low-income areas to attend the camp. However, this will make future data collection difficult if working with both the Eugene and Springfield School districts, and requires some consideration.

By the time another study can be performed, the Springfield School District should have a fully functioning database, including information on whether or not students enter college, which will allow for the actual differences between the students' choices to be evaluated. As there are missing observations from both the younger and older cohort because of students who were missing information for key variables, the results of this study are biased. The missing

observations tend to come from those who were troubled in school, suggesting our results are biased upwards. With more observations in the future, the study should have more power, leading to a more definitive answer of the effects of the SAIL program.

References

ACT's Activity Publication. (2011). "College Retention Rates Improving at Two-Year Schools, Declining at Four-Year School". Accessed June 12. 2012. Retrieved from http://www.act.org/activity/spring2011/retention.html

Angrist, J., Lang, D., & Oreopoulos, P. (2009). "Incentives and services for college achievement: Evidence from a randomized trial". *American Economic Journal: Applied Economic*s. 1(1), 136–163.

Caliendo M., Kopeinig S. (2008). "Some Practical Guidance for the Implementation of Propensity Score Matching". *Journal of Economic Surveys*. 22(1), 31-72.

Dobbie, W. and Fryer, R. (2011). "Are High-Quality Schools Enough to Increase Achievement Among the Poor? Evidence from the Harlem Children's Zone". *American Economic Journal: Applied Economics* 3(3): 158-187.

Dynarski, S. (2005). "Building the Stock of College-Educated Labor". *National Bureau of Economic Research*.

European Commission. (2009, November 24). "Propensity score matching." Accessed March 17, 2012. Retrieved from http://ec.europa.eu/regional_policy/sources/docgener/evaluation/evalsed/sourcebooks

Geiser, S., & Santelices, V. (2004). "The role of Advanced Placement and honors courses in college admission". *Center for Studies in Higher Education*. University of California, Berkeley.

Gullat, Y., & Jan, W. (2002). "How do pre-collegiate academic outreach programs impact college-going among underrepresented students?" *Boston: Pathways to College Network.*

Heckman, J. J., Masterov, D.V. (2004). *The productivity argument for investing in young children* (Working Paper No. 5). Retrieved from http://jenni.uchicago.edu/Invest/.

Hullsiek, K. H., & Louis, T. A. (2002). "Propensity score modeling strategies for the causal analysis of observational data". *Biostatistics*, 2(4), 179–193.

Malecki, C. K., & Demaray, M. K. (2006). "Social support as a buffer in the relationship between socioeconomic status and academic performance". *School Psychology Quarterly*, 21: (375-395).

Moretti, E. (2004). "Estimating the Social Return to Higher Education: Evidence from Longitudinal and Cross-Sectional Data". *Journal of Econometrics*, 121: (157-212).

Myers, D. et al. (2004). *The Impacts of Regular Upward Bound: Results from the Third Follow Up Data Collection.* Washington, DC: U.S. Department of Education.

Palmer, S. (2012, January 27). "Fewer graduate at several schools: five Eugene and Springfield high schools see a drop in the graduation rate". *The Register Guard*.

Peikes, D., Moreno, L., & Orzol, S. (2008). "Propensity score matching: A note of caution for evaluators of social programs". *The American Statistician*, *62*(3), 222-231. Retrieved from http://dx.doi.org/10.1198/000313008X332016

Reardon, S. (2011). "The Widening Achievement Gap Between the Rich and the Poor: New Evidence and Possible Explanations". *Whither Opportunity? Rising Inequality, Schools and Children's Life Chances:* (91-116.) New York: Russell Sage Foundation.

Rosenbaum, P.R., & Rubin D.B. (1985). "Constructing a control group using multivariate matched sampling methods that incorporate the propensity score". *The American Statistician* 39(1): 33-38.

Rumberger, R. W. (2010). "Education and the reproduction of economic inequality in the United States: An empirical investigation". *Economics of Education Review* 29(2): 246–254.

University of Oregon (2011). *SAIL: Summer Academy to Inspire Learning*. Retrieved 25 April 2012 from University of Oregon: http://sail.uoregon.edu/about.html

U.S. Census. (2008). *Educational Attainment in the United States: 2008. Detailed Tables*. http://www.census.gov/population/www/socdemo/education/cps2008.html

Wolfe, B., & Haveman R. (2001). "Accounting for the Social and Non-Market Benefits of Education". *The Contribution of Human and Social Capital to Sustained Economic Growth and Well-Being*. Human Resources Development Canada and Organisation for Economic Co-operation and Development.

Zhao, E. (2011, October 20) "High School Dropout Rates for Minority and Poor Students Disproportionately High". *Huffington Post*.

Appendix

A. Model Specification

- Model used to develop propensity score:

$$Sail = \beta_1 + \beta_2 Sex_i + \beta_3 Free\ Lunch_i + \beta_4 Reduced\ Lunch_i + \beta_5 OakReading_i + \beta_6 ESL_i + \beta_7 OaksReading_i + \beta_8 Race - White_i + \beta_9 Race - Hispanic_i + \beta_{10} Special\ Education_i + u_i,$$

- Older cohort models (Includes all variables used. Not all regressions used every variable listed):

$$Algebra2 = \beta_1 + \beta_2 Sex_i + \beta_3 Free\ Lunch_i + \beta_4 Reduced\ Lunch_i + \beta_5 OakReading_i + \beta_6 ESL_i + \beta_7 OaksReading_i + \beta_8 Race - White_i + \beta_9 Race - Hispanic_i + \beta_{10} OaksMath_i + \beta_{11} Attendance\%_i + u_i,$$

$$AP = \beta_1 + \beta_2 Sex_i + \beta_3 Free\ Lunch_i + \beta_4 Reduced\ Lunch_i + \beta_5 OakReading_i + \beta_6 ESL_i + \beta_7 OaksReading_i + \beta_8 Race - White_i + \beta_9 Race - Hispanic_i + \beta_{10} OaksMath_i + \beta_{11} Attendance\%_i + u_i,$$

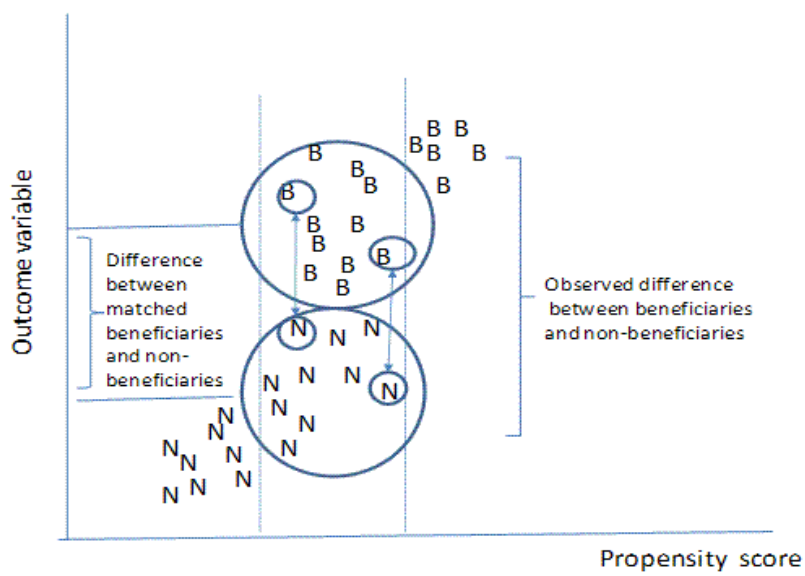
$$GPA = \beta_1 + \beta_2 Sex_i + \beta_3 Free\ Lunch_i + \beta_4 Reduced\ Lunch_i + \beta_5 OakReading_i + \beta_6 ESL_i + \beta_7 OaksReading_i + \beta_8 Race - White_i + \beta_9 Race - Hispanic_i + \beta_{10} OaksMath_i + \beta_{11} Attendance\%_i + u_i,$$

- Younger cohort model (Includes all variables used. Not all regressions used every variable listed):

$$GPA_i = \beta_1 + \beta_2 SAIL_i + \beta_3 RejectSAIL_i + \beta_4 NotRecommended_i + \beta_5 OakReading_i + \beta_6 FreeLunch_i + \beta_7 ReducedLunch_i + \beta_8 Race - White_i + \beta_9 Race - Hispanic_i + \beta_{10} OAKSMath_i + \beta_{11} Attendance\%_i + \beta_{12} MinorIncidents_i + \beta_{12} MajorIncidents_i + u_i$$

B. ("Propensity score matching" 24/11/2009)

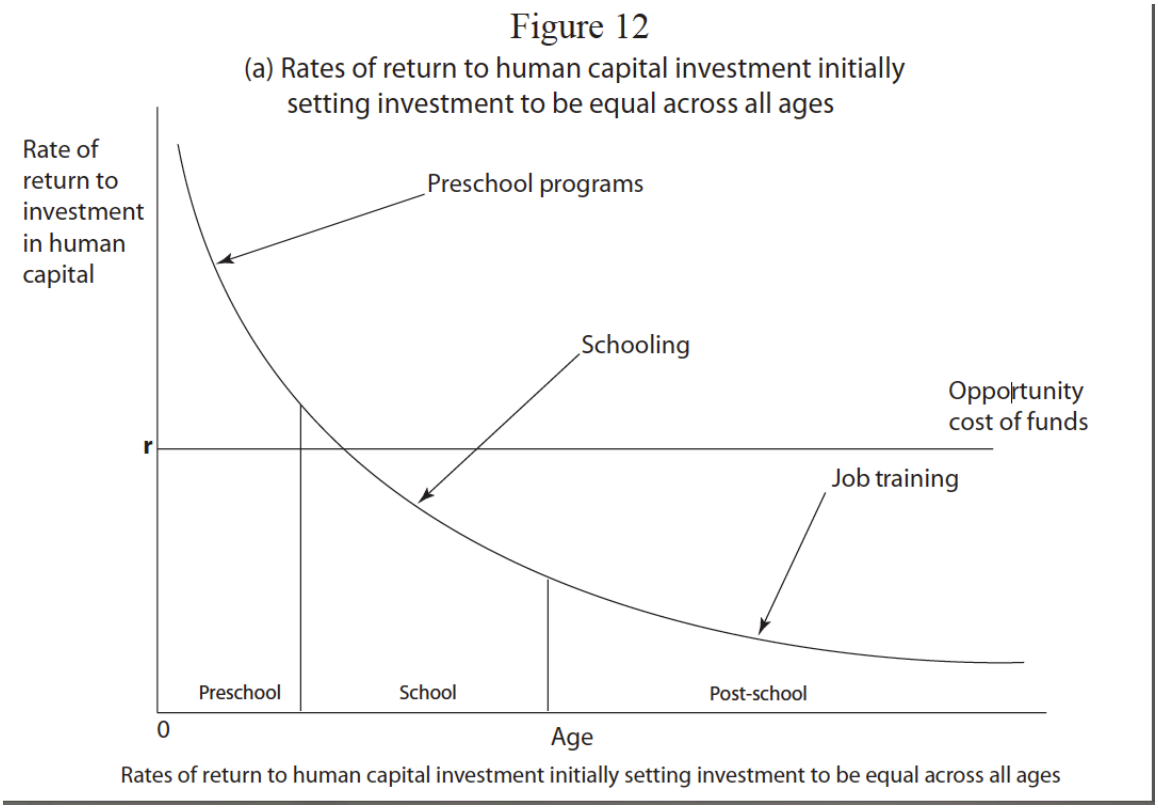
Figure 1. A graphical representation of matching on the propensity score



This graphic depiction attempts to explain the mechanical function of propensity score matching. If one thinks of "outcome variable" as the y variable and "propensity score" as the x variable on a typical Cartesian plane, then the circle with the greater y value is the group most likely to

receive treatment. Within each circle, there are various subjects with differing characteristics. If one were to add a scale to the graph and circle an arbitrary subject from each group, the vertical distance between the two is the measured observable difference between the subjects. Their respective x value (or propensity score value) determines the probability of receiving the intervention. This graph is meant to convey that ideally, the distance between B and N should be minimized.

C. Chart of return to human capital investments (Heckman, 2004)



The above chart created by Heckman, details the negative correlation between age and the return rate to investment in human capital. While the line **r**, which signifies the quantity of return to the investment, is arbitrarily drawn, it is meant to impress upon the reader that after a given amount of investment, the opportunity cost of investing a person outweighs the benefit he or she might bestow on the economy.

D. Additional data tables and their explanations

Listed below are the results of the SAIL intervention with Robust S.E. and Bootstrapped S.E.

Table 6A:Effect of SAIL Intervention on College Readiness using Robust S.E

Dependent Variable	Algebra 2			AP			GPA		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Independent Variable									
ESL	0.5031093	0.5930412	0.5530248	6.671022	6.896693	1.876537	0.5946374	0.6605178	0.5837874
	(0.6608056)	(0.6445161)	(0.6283454)	(439.3107)	(450.3624)	(0.8220062)**	(0.2437901)**	(0.3233647)**	(0.4788439)
SAIL-Attended	0.350119	0.3939148	0.3612061	-0.0970071	-0.0196913	-0.030861	0.1053381	0.1788603	0.1585715
	(0.3158158)	(0.3119675)	(0.3087801)	(0.3652715)	(0.3470508)	(0.3215732)	(0.1471043)	(0.1733755)	(0.1725036)
8th Grade OAKS-Reading	0.0693166	0.0704421	0.0644955	0.1248196	0.1177899	0.0837363	0.0233064	0.0281814	0.0225118
	(.0319477)**	(0.0316881)**	(0.0309973)**	(0.0405142)***	(0.0383205)***	(0.0334445)**	(0.0106654)**	(0.0119562)**	(0.0124562)*
8th Grade OAKS-Math	-0.0070306	-0.0068831	-0.0063794	0.0027956	0.0024958	0.0009146	0.0055791	0.0063791	0.0061438
	(0.0185428)	(0.0182874)	(0.0182079)	(.0201697)	(.0195462)	(0.0186982)	(0.0075084)	(0.009271)	(0.0097825)
Male	-0.4978593	-0.4726298	-0.4624025	-0.6044086	-0.5834952	-0.3563795	-0.0624697	-0.0757635	-0.0697049
	(0.2989324)*	(0.2930972)	(0.2914563)	(0.3511539)*	(0.3330057)*	(0.3059605)	(0.1424334)	(0.1708836)	(0.1778694)
Race-Hispanic	0.7038296	0.6972829		10.07895	11.03114		0.8222737	0.8470163	
	(0.6353214)	(0.6313672)		(648.1665)	(664.3669)		(0.2545764)***	(0.2828636)***	
Race-White	0.5223912	0.471444	-0.0412123	9.736626	10.69394	0.1856435	0.9907826	0.8999999	0.2956551
	(0.5772229)	(0.5801045)	(0.3338656)	(648.1666)	(664.3671)	(0.3490207)	(0.229835)***	(0.2440716)***	(0.183522)
Lunch-Free	-0.3374542	-0.3324709	-0.3032773	-0.7595128	-0.6966218	-0.468055	0.2161759	0.1973549	0.244588
	(0.3204796)	(0.3135782)	(0.30997)	(0.3770143)**	(0.3550387)**	(0.3242752)	(0.1481323)	(0.1779638)	(0.1768646)
Lunch-Reduced	-0.7657299	-1.008492	-1.072539	-0.1048434	-0.6297935	-0.7877829	-0.1243744	-0.3473413	-0.437359
	(0.6351628)	(0.6430375)	(0.6330623)*	(0.7109844)	(0.7109585)	0.6576127	(0.3337977)	(0.2981012)	(0.3153307)
Attendance%	7.140331			12.70052			8.516875		
	(4.337394)*			(5.266122)**			(1.903544)***		
Constant	-22.2863	-15.67304	-13.85734	-52.25645	-39.20922	-20.3481	-13.16021	-6.245443	-4.236234
	(7.644302)***	(6.262205)**	(5.964671)**	648.2466	(664.4144)	(6.651077)***	(2.637182)***	(2.722335)**	(2.826468)
Number of Observations	93	93	93	93	93	93	93	93	93
LR Chi(Independnt Variable	18.34	15.99	14.72	42.05	34.82	20.15			
Prob>Chi2	0.0495	0.0671	0.0648	0	0.0001	0.0098			
Pseudo R2	0.15	0.1298	0.1195	0.3512	0.2908	0.1683			
F-Statistic							7.42	4.77	2.78
Prob > F							0	0	0.0088
R-squared							0.399	0.2062	0.1484
Root MSE							0.66507	0.75971	0.78219

*=P<.10, **=P<.05, ***=P<.01

Table 6B: Effect of SAIL on College Readiness using Bootstrapped S.E.

Dependent Variable	Algebra 2			AP			GPA		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Independent Variable									
ESL	0.5028206	0.5930412	0.5530248	6.671022	6.896693	1.876537	0.5946374	0.6605178	0.5837874
	(0.7585605)	(0.9965496)	(1.042908)	(1.131483)***	(0.6275277)***	(0.9273648)**	(0.3259857)*	(0.5462986)	(0.52601)
SAIL-Attended	0.3502374	0.3939148	0.3612061	-0.0970071	-0.0196913	-0.030861	0.1053381	0.1788603	0.1585715
	(0.3718416)	(0.3424641)	(0.3095372)	(0.2850902)	(0.3812572)	(0.3342295)	(0.1564311)	(0.1981314)	(0.1914109)
8th Grade OAKS-Reading	0.0693392	0.0704421	0.0644955	0.1248196	0.1177899	0.0837363	0.0233064	0.0281814	0.0225118
	(0.0281788)**	(0.0380722)*	(0.046203)	(0.0267539)***	(0.0429998)***	(0.0332323)**	(0.0136607)*	(0.0139819)**	(0.0127644)*
8th Grade OAKS-Math	-0.0070119	-0.0068831	-0.0063794	0.0027956	0.0024958	0.0009146	0.0055791	0.0063791	0.0061438
	(0.0228086)	(0.019528)	(0.0219113)	(0.0139431)	(0.0186355)	(0.0258514)	(0.0072151)	(0.009218)	(0.0091613)
Male	-0.4987759	-0.4726298	-0.4624025	-0.6044086	-0.5834952	-0.3563795	-0.0624697	-0.0757635	-0.0697049
	(0.2769771)*	(0.4137507)	(0.338794)	(0.3489437)*	(0.4226487)	(0.3662844)	(0.1395591)	(0.1708836)	(0.2051415)
Race-Hispanic	0.7046346	0.6972828		10.07895	11.03114		0.8222737	0.8470163	
	(2.098438)	(1.415837)		(0.8657837)***	(0.7207167)***		(0.2895039)***	(0.3990877)**	
Race-White	0.5224541	0.471444	-0.0412123	9.736626	10.69394	0.1856435	0.9907826	0.8999999	0.2956551
	(2.018821)	(1.37798)	(0.420126)	(1.2225)***	(0.9081279)***	(0.5069199)	(0.2727768)***	(0.3026281)***	(0.1874242)
Lunch-Free	-0.3385392	-0.3324709	-0.3032773	-0.7595128	-0.6966218	-0.468055	0.2161759	0.1973549	0.244588
	(0.3531086)	(0.3939757)	(0.4607805)	(0.4795691)	(0.4323338)	(0.4221896)	(0.1419336)	(0.1701535)	(0.1849302)
Lunch-Reduced	-0.7636582	-1.008492	-1.072539	-0.1048434	-0.6297935	-0.7877829	-0.1243744	-0.3473413	-0.437359
	(0.6756828)	(0.4080922)**	(0.5328126)**	(0.5731577)	(0.5116225)	(0.470715)*	(0.4330997)	(0.30347)	(0.3255661)
Attendance%	7.205661			12.70052			8.516875		
	(3.674176)**			(3.99709)***			(2.258777)***		
Constant	-22.35923	-15.67304	-13.85734	-52.25645	-39.20922	-20.3481	-13.16021	-6.245443	-4.236234
	(8.106363)***	(9.667025)	(9.360008)	(7.991754)	(8.768905)***	(8.684668)***	(3.116768)***	(3.027572)**	(2.818078)
Number of Observations	93	93	93	93	93	93	93	93	93
Replications	28	26	25	13	18	20	49	47	47
Wald chi2	20.42	21.81	8.07	76539.32	2888.33	49.59	44.81	27.6	26.97
Prob>chi2	0.0255	0.0095	0.427	0	0	0	0	0.0011	0.0007
Pseudo R2	0.1565	0.1298	0.1195	0.3512	0.2908	0.1683			
R-squared							0.399	0.2062	0.1484
Adjusted R-squared							0.3257	0.1201	0.0673
Root MSE							0.6651	0.7597	0.7822

*=P<.10, **=P<.05, ***=P<.01

The bootstrapped results differ little from the robust standard error results. This may occur because our sample is reflective of the population (except for slight deviations in race, which are explained in the data description) and, according to our differences in means table, randomized. One noteworthy difference is the increase in significance of being White across models in the Bootstrapped dataset relative to Robust S.E. This effect is most likely capturing the over representation of Whites in the dataset.

The data table below is of the probit models used to determine SAIL's effect on college readiness using bootstrapped standard errors. This data table was constructed for comparing robust standard errors with bootstrapped standard errors. For the marginal effects using bootstrapped standard errors, see pages 26-27.

Table 7: Marginal Effects of SAIL on College Readiness using Robust S.E.

Dependent Variable	Algebra 2			AP Courses?		
	(1)	(2)	(3)	(1)	(2)	(3)
Independent Variable						
Free Lunch	-0.1118821	-0.1221188	-0.1234519	-0.1634504	-0.1596621	-0.1747018
	(0.11013)	(0.11156)	(0.10959)	(0.10664)	(0.07955)**	(0.08109)**
Reduced Lunch	-0.3037348	-0.2902621	-0.2349906	-0.2276078	-0.1147653	-0.0245174
	(0.11918)**	(0.12696)***	(0.15836)***	(0.1367)*	(0.08756)	(0.18072)
Race-Hispanic		0.2688255	0.2708478		0.9996253	0.999205
		(0.24931)	(0.24136)		(0.00028)***	(0.00062)***
Race-White	-0.0154406	0.1672898	0.1826705	0.0651632	0.9824042	0.966862
	(0.12148)	(0.20011)	(0.18652)	(0.12368)	(0.00842)***	(0.01807)***
Male	-0.1703059	-0.1735189	-0.1816693	-0.1262908	-0.1383574	-0.1443909
	(0.10641)	(0.10756)*	(0.10605)*	(0.10693)	(0.07677)*	(0.07995)*
ESL	0.2165613	0.23206	0.19611	0.6157282	0.9023707	0.8992229
	(0.36927)	(0.3773)	(0.39751)	(0.15822)***	(0.02571)***	(0.02858)***
SAIL Attendance	0.136812	0.1489688	0.1316083	-0.0110254	-0.004757	-0.0233477
	(0.11281)	(0.11373)	(0.11116)	(0.11153)	(0.08074)	(0.0806)
OAKS-Reading	.0240906	0.0262427	0.0256647	0.029984	0.0285431	0.0304984
	(0.01129)**	(0.0116)	(0.01186)**	(0.01119)***	(0.00953)***	(0.01114)***
OAKS-Math	-.0023829	-0.0025643	-0.0025953	0.0003275	0.0006048	0.0006831
	(0.00681)	(0.0068)	(0.00656)	(0.00681)	(0.00501)	(0.00487)
Attendance			2.667046			3.103238
			(1.53091)*			(1.43303)**

*=P<0.10, **=P<0.05, ***=P<0.01

The above data table is for comparing the robust standard errors and bootstrapped standard errors of the probit models' marginal effects. Interestingly, magnitude of the bootstrapped standard errors is smaller, implying that the robust command from Stata yielded biased standard errors. Since bootstrapping forces the sample standard errors to better match those of the population, these standard errors give more accurate t-test values than the robust standard errors.

E. Original data for the Older Cohort, post-graduation outcome by graduation year.

Table 8: Older Cohort: College Outcomes for SAIL Students

Graduation Year	High School	4-year	Other post-secondary	Military	Other
repeating	Willamette				Repeat
2012	SHS		LCC		
2012	SHS		LCC		
2012	A-3		LCC		
2012	A-3		LCC		
2012	Sheldon		Chemekata		
2012	Sheldon		LCC		

2012	SHS			Army	
2012	SHS		LCC		
2012	SHS			Military	
2012	South				Graduated
2012	Willamette	UO			
2012	SHS			Military	
2012	Sheldon		LCC		
2012	Sheldon		LCC		
2012	SHS				Graduated
2012	A-3 before moving		Palamar CC		
2012	SHS		Beauty		
2012	SHS	UO			
2012	SHS		LCC		
2012	SHS		LCC		
2012	SHS	UO			
2012	SHS	UO			
2012	SHS		LCC		
2012	SHS	UO			
2012	Willamette				Graduated
2011	SHS				Graduated
2011	SHS	OSU			
2011	SHS	UO			
2011	SHS		LCC		
2011	SHS	OSU			
2011	SHS		LCC		
2011	SHS			Military	
2011	SHS	UO			
2011	SHS				Unknown
2011	SHS			Military	
2011	SHS		Central OR CC		
2011	SHS	George Fox			
2011	Moved				Unknown
2011	SHS		LCC		
2011	SHS	UO	LCC		
2011	SHS				Unknown
2011	SHS				Unknown
2011	Willamette	UO			
2011	South	UO			
2011	South		LCC		
2010	SHS		Blue Mt. CC		
2010	SHS		LCC		
2010	SHS		LCC		
2010	SHS	UO			
2010	SHS	UO			
2010	SHS	UO			
2010	SHS		LCC		
2010	SHS		LCC		
2010	SHS		Pioneer Pacific		
2010	Willamette				Unknown

2010	SHS	George Fox			
2010	SHS				Unknown
2010	Moved				Unknown
2010	Did not grad?				Unknown
2010	SHS		LCC		
2010	Thurston				Unknown
	Totals	**25.81%**	**41.94%**	**8.20%**	**17.74%**