Isolating the Impact of Learning Communities and First-Year Residence Halls on First-Year Student Retention and Success

Robert C. Guell
Project Manager of the Lilly Project
to Transform the First-Year Experience
Indiana State University

Abstract

This paper showcases the success of Indiana State University's *Lilly Project to Transform the First-Year Experience* in increasing first-year retention and GPA attainment and provides statistically sound and significant evidence that Learning Communities and First-Year Residence Halls played a substantial role in that success. By employing logistical and linear regression techniques, we know that one-quarter of the eight-percentage point increase in retention since 1994 is due to learning communities and that Learning Community participants achieve higher first-semester GPA's. We know that one-fifth of the increase is due to participation in first-year residence halls and that such participation also increases first-semester GPA's.

Introduction

Indiana State University's *Lilly Project to Transform the First-Year Experience* began in 1997 with a 5-year $2 million grant from the Lilly Endowment. The Endowment's goal was to increase the educational attainment of Indiana adults and to accomplish this it funded a variety of efforts in the state's public and private universities. ISU's strategy to raise graduation rates was to increase first-year retention rates by employing first-year Learning Communities. Simultaneously, ISU's Residential Life division of its Office of Student Affairs was reforming the housing arrangements for first-year students. Starting from the point in 1994 when the grant application was made, first-year retention rates at ISU have risen from 64% to 72%. This paper reports the results of statistical analysis that isolates the marginal contribution of these reforms to that eight-point increase.

Using data from 1999 through 2001 and employing linear and logistic regression techniques it was determined that much of the eight-percentage point increase can be attributed directly to these programmatic endeavors. Specifically, using logistical regression, it was shown that one-quarter of the increase is due to Learning Communities (LCs) alone. Similarly, those same logistical regressions indicated that the creation of special first-year residence halls (FYRH) accounted for one-fifth of the increase. It was also determined using linear regression that these programs increased first semester GPA's in a statistically and practically significant fashion. LC participants had GPA's approximately .16 points higher than their equally situated non-LC colleagues and FYRH program participants had GPA's .11 points higher.

Finally to gauge the financial impact of these reforms we noted that the eight-percentage point increase in first-year retention not only translated into additional sophomores, but more juniors and seniors as well. Accounting for the increases in the numbers of students at all levels that result from these retention increases it was determined that ISU, with nearly 12,000 students, bills $2.1 million more in tuition annually than it otherwise would have. Considering that these programs cost less than a fifth of that, these reforms are obviously working quite well from a financial standpoint as well.

Data

The data for this study comes from the academic files of approximately 6,000 first-year students who enrolled at ISU from 1999 to 2001. The variables used for predicting retention and GPA attainment are fairly standard and are broken into two categories: pre-entry variables and programmatic variables. Pre-entry variables are important for two reasons: better students do better than poorer ones and better students are more likely to be attracted to curricular innovations than poorer ones. It is for these reasons that verbal and math SAT scores, high school GPA, and class rank, whether or not the student is "first-generation" and the income class of the students' families are included.

Also included in the pre-entry variable list is an indicator variable, CORE40, which accounts for the fact that the State of Indiana created a state-wide curriculum for its college-bound high school students. This "Core 40" curriculum, as the name suggests, requires forty credit hours with an eye toward college preparation. Among the important features, the curriculum requires 3 years of math and science, 4 years of Language Arts as well as two years of foreign language. Entrance into Indiana's premier Research 1 institutions (Indiana University and Purdue University) requires such a "Core40" diploma while other institutions merely encourage prospective students to work on this college track. Approximately 60% of ISU's students receive such a high school diploma. The presumption is that students on a college track in high school will perform better in college once they get there than students not on a college track.

The remaining variables are the focus of this paper. Participation in Learning Communities, the centerpiece of the Lilly Project, is voluntary for many and mandatory for few. Majors in the Schools of Technology, Nursing and Business have LCs that are mandatory unless students are transferring from other institutions or have scheduling difficulties that preclude participation. Majors in Pre-Med, Communication, Social Work, Legal Studies, Art, Music, Criminology, Psychology, or those who are "Open Preference" or conditionally admitted may chose to participate.

The final variable indicates participation in the first-year residence hall program. Started in the year prior to LCs (1997) this program was not funded by the Lilly Project but was established by the Office of Student Affairs and its Residential Life division to enhance retention using what was known about the importance of social adjustment. The program included three important features: the concentrating of most freshman in three halls, the establishment of Academic Peer Advocates (upper-class students each assigned to two floors with the task of helping student academic adjustment) and the concentration of programming in the first weeks of the semester to combat general problems of adjustment, as well as specific ones related to alcohol and drug use and sexual health.

Because Core40 is only an appropriate variable for students who graduated from high school in the last 3 years and because we conduct the survey of students where they give us their parents' income and educational attainment at our summer registration and advising program, including these variables impacts the size of the dataset. Thus, including the Core40 variable alone reduces the dataset from 3014 observations to 2863,

while including the income and first-generation variables alone reduces the dataset to 2727. Including both variables reduces the dataset to 2605.

Variable	**Description**
SATM	SAT Math (or ACT equivalent)
SATV	SAT Verbal (or ACT equivalent)
HSGPA	High School GPA
HSRank	High School Class Rank
FG	First Generation College Student =1 if yes =0 if no
$20K-$40K	Family income between $20,000 and $40,000
$40K-$60K	Family income between $40,000 and $60,000
>$60K	Family income greater than $60,000
Core40	College Track High School Diploma (only relevant for Indiana High School Students =1 if yes =0 if no
LC	Learning Community Participation =1 if yes =0 if no
FYRH	First-Year Residence Hall =1 if yes =0 if no

Methodology-Determining The Impact of Programs

The marginal impacts of Learning Communities and First-Year Residence Halls were determined using linear and logistic regression. Because the dependent variable in a retention equation is a binary variable, linear regression, while yielding parameters that are easy to interpret, is inappropriate. In such a case, logistical regression is used to estimate the degree to which any particular variable impacts retention. It is important to recall that both techniques create estimates that control for the impact of other variables. In linear regression the parameters suggest the impact of variables directly while in logistical regression the impact of one independent variable on the dependent variable depends on the level of the other independent variables.

For the present analysis logistical regression is appropriate to estimate the factors involved in retention. To determine the marginal impact of the particular program requires that the parameter estimates from the logistical regression be used to create estimated retention probabilities both with and without the programs' parameters. Thus if the underlying logistical regression is

$$Y = b_{pe} X_{pe} + b_{LC} P_{LC} + b_{FYRH} P_{FYRH}$$

where

Y	=1 if retained, =0 otherwise,
X_{pe}	a vector of pre-entry variables,
P_{LC}	=1 if involved in Learning Communities, =0 otherwise,
P_{FYRH}	=1 if involved in First-Year Residence Halls, =0 otherwise,

then the marginal impact of the program is the difference in the predicted probabilities.

Marginal Impact

The marginal impact of the program is taken by getting the probability of being retained given participation in the program(s), and comparing that to the probability of being retained given non-participation in the program(s). For example, the marginal impact of LCs is

$$MI = \hat{Y}_{with} - \hat{Y}_{without}$$

where

$$\hat{Y}_{with} = \frac{e^{\hat{b}_{pe} X_{pe} + \hat{b}_{LC} P_{LC}}}{1 + e^{\hat{b}_{pe} X_{pe} + \hat{b}_{LC} P_{LC}}}$$

$$\hat{Y}_{without} = \frac{e^{\hat{b}_{pe} X_{pe}}}{1 + e^{\hat{b}_{pe} X_{pe}}}$$

The marginal impact of the First-Year Residence Hall program is similarly established.

Marginal Impact of Both Programs

Unlike linear regression where the marginal impacts of each program can be directly aggregated to generate the impact of all programs combined, in logistical regression the sum of the marginal impacts can be greater than or less than the marginal impact of participating in all of the programs. Generating the marginal impact of all programs combined on retention is performed in a similar fashion to that shown above and is

$$MI = \hat{Y}_{all} - \hat{Y}_{none}$$

where

$$\hat{Y}_{all} = \frac{e^{b_{pe} X_{pe} + b_{LC} P_{LC} + b_{FYRH} P_{FYRH}}}{1 + e^{b_{pe} X_{pe} + b_{LC} P_{LC} + b_{FYRH} P_{FYRH}}}$$

$$\hat{Y}_{none} = \frac{e^{\hat{b}_{pe} X_{pe}}}{1 + e^{\hat{b}_{pe} X_{pe}}}$$

The Effect on GPA

Estimating the impact these programs have on first semester GPA attainment is substantially less difficult. Using linear regression[1] the parameter estimates are directly read from the equation

[1] while a two-limit Tobit would be more appropriate in this circumstance there are so few perfect 4.0 GPAs and (imperfect) 0.0 GPAs to make such an accommodation to statistical niceties practically unimportant.

$$GPA = g_{pe} X_{pe} + g_{LC} P_{LC} + g_{FYRH} P_{FYRH}$$

and the marginal impacts of LCs and participation in FYRHs are the parameters ?$_{LC}$ and ?$_{FYRH}$ respectively.

Results- Determining The Impact of Programs

Though flawed statistically, linear regression results for determining the impact of programs on retention are useful in that they can give first-pass estimates in an easily understood fashion. These results are shown in Table 1 and suggest that participation in Learning Communities and First-Year Residence Halls significantly increase retention. The results from Table 2 suggest a similar set of conclusions using logistical regression. Table 3 displays the marginal impacts of these programs.

Learning Communities and First-Year Residence Halls each independently contribute to retention and together account for just under half of the overall increase in retention since they were employed. Depending on the alternate specifications, participation in Learning Communities increases retention by 4.8 to 5.8 percentage points. Similarly, participation in First-Year Residence Halls increases retention by 3.3 to 3.5 percentage points.

Because logistical regression parameters, unlike linear regression parameters, cannot be simply added together to generate combined effects of both programs, Table 3 also reports the aggregate impact for participation in both programs. Students who choose to participate in both programs are retained at between 7.7% and 8.7% higher rates than their equally academically situated first-semester colleagues.

Learning Communities and First-Year Residence Halls also contribute to first-semester GPA attainment in a statistically and practically important way. As is shown in Table 4, across differing specifications participation in LCs increases first-semester GPA attainment between .155 and .165 points. FYRH participation increases GPA attainment between .108 and .125 points. As a result, we can say that participation in this programs leads to not only higher retention rates but to GPAs that are .263 to .290 points higher than that attained by their equally academically situated first-semester colleagues.

Methodology - Estimating Financial Impacts

An eight percentage point increase in the first-year retention rate of any University has enormous financial consequences but those consequences partly depend on whether those students are retained until graduation, retained through graduation in roughly the same fashion as those who were retained previously, or drop out soon after their sophomore year. At ISU our second and third year retention rates rose and our conditional retention rates[2] remained relatively constant. As a result, the increase in first-year retention continues to have ripple effects as there are not only more sophomores each year but more juniors and seniors down the road. Estimating the magnitude of the tuition

[2] For instance, the conditional second-year retention rate is the percentage of sophomores who become juniors.

generated from these effects and establishing the programs to which these savings can be attributed is the purpose of this section.

As can be seen in Table 5, with an eight-percentage point increase in first-year retention, there are 164 additional sophomores from a freshman class of 2050. Because the conditional second, third and fourth-year retention rates remained roughly constant (at 81.9%, 84.7%, 44%[3],) ISU gains 134.4 additional juniors and 164 additional fourth and fifth year seniors. Therefore the long-term impact of the eight-point increase in retention results in a total of 462.4 more students enrolling at ISU. With in-state students making up 85% of the student body and paying annual tuition of $3,744 and out-of-state students making up the remaining 15% and paying tuition of $9,346, ISU nets $2,119,729 in additional billable tuition dollars.[4]

Table 4 serves to parse the responsibility for that $2.1 million to the programs that increased first-year retention in the first place. To establish these amounts, the marginal impact figures from Table 3 are used to determine the percentage of the eight point increase in retention attributable to each program. Because Learning Communities increase the first-year retention rate for participants by approximately 5.4 percentage points and because 40% of first-year students enroll in those LCs, this innovation is responsible for 26.5% of the increase in retention and $561,728 in increased tuition. Learning Communities alone cost less than $100,000 annually to operate.

Similarly, because First-Year Residence Halls house 52% of first-year students and because being in a FYRH adds 3.4 percentage points to the retention rate of those participating students, the program can accurately claim 21.5% of the credit for that eight percentage point increase in first-year retention. This amounts to $468,460 in additional tuition revenue. The additional cost to the campus for the Academic Peer Advocates and the programs surrounding First-Year Residence Halls is also less than $100,000.

While there is overhead in these first-year programs, much of that overhead goes for other programs aimed at retention that are more difficult to measure. For instance, establishing precisely how much a reform to a currently existing program impacts retention is a more difficult task. For instance students who participate in our reformed Sycamore Advantage summer registration program get a benefit. On the other hand a summer registration program existed prior to the reforms. Similarly, more than $600,000 was devoted over the last five years to faculty and staff development around first-year issues. Establishing how much of the unaccounted for 52% of the eight percentage point increase is attributable to these efforts will be the subject of further study.

---

[3] ISU's four-year graduation rate is 22%. Thus 44% of 4th year seniors become 5th year seniors, 45% graduate and therefore do not return and the rest either complete their degree after six or more years of do not finish their degree at all.

[4] The degree to which this additional money is forgone in higher financial aid grants from the institution is not obvious. The financial aid budget is a set amount and not a direct function of total billable tuition. Using that logic the entirety of the $2.1 million is a net increase. On the other hand, the history of financial aid needs drive the financial aid budget.

Conclusion

The Indiana State University's effort to increase first-year retention has clearly been successful. The eight-point increase in first-year retention has generated a number of important benefits. The most significant of these are the 464.5 additional students and $2.1 million increase in tuition revenue. This paper has shown that just under half of these benefits are directly attributable to two reforms, Learning Communities and First-Year Residence Halls. The paper also demonstrates that these reforms impact student GPA attainment as well.

Table 1
Fall-to-Fall Retention
Linear Regression
Fall 1999 and 2000 Co-horts

	Specification 1		Specification 2		Specification 3		Specification 4	
N	2727		3014		2605		2863	
	Para. Est	Sig	Para. Est	Sig	Para. Est	Sig	Para. Est	Sig
Intercept	0.500	***	0.471	***	0.516	***	0.470	***
LC	0.045	***	0.044	***	0.054	***	0.055	***
SATM	-0.00004		-0.00001		-0.000050		-0.00003	
SATV	-0.00032	***	-0.00028	**	-0.000327	***	-0.00029	**
HSGPA	0.122	***	0.122	***	0.111	***	0.117	***
HSRank	0.001		0.001		0.001		0.001	
FYRH	0.032	**	0.030	**	0.033	**	0.031	**
FG	-0.016				-0.023			
Core40					0.020		0.029	*
$20K-$40K	0.016				0.019			
$40K-$60K	0.032				0.025			
>$60K	0.043	*			0.037			
R^2	0.031		0.031		0.034		0.035	
adj R^2	0.028		0.029		0.030		0.033	

Table 2
Fall-to-Fall Retention
Logistic Regression
Fall 1999 and 2000 Co-horts

	Specification 1	**Specification 2**	**Specification 3**	**Specification 4**
N	2727	3014	2605	2863
	Para. Est Sig	Para. Est Sig	Para. Est Sig	Para. Est Sig
Intercept	-0.412	-0.548	-0.340	-0.565
LC	0.269***	0.261***	0.329***	0.324***
SATM	-0.00014	0.00009	-0.00019	-0.00008
SATV	-0.00182***	-0.00160**	-0.00185**	-0.00158**
HSGPA	0.736***	0.716***	0.681***	0.700***
HSRank	0.003	0.003	0.005	0.004
FYRH	0.189**	0.176**	0.194**	0.178**
FG	-0.098		-0.135	
Core40			0.107	0.150*
$20K-$40K	0.087		0.101	
$40K-$60K	0.168		0.128	
>$60K	0.237*		0.202	

Table 3
Marginal Impacts of Programmatic Variables
From the Logistic Regression Models

	Specification 1	Specification 2	Specification 3	Specification 4
FYRH	3.4%	3.3%	3.5%	3.3%
LC	4.8%	4.8%	5.7%	5.8%
Aggregate	7.7%	7.7%	8.7%	8.6%

Table 4
First-Semester Grade Point Average
Linear Regression
1999-2001

	Specification 1		**Specification 2**		**Specification 3**		**Specification 4**	
N	2727		3014		2605		2863	
	Para. Est	Sig	Para. Est	Sig	Para. Est	Sig	Para. Est	Sig
Intercept	0.503	***	0.391	***	0.494	***	0.376	***
LC	0.155	***	0.162	***	0.157	***	0.165	***
SATM	0.00002		0.00006		-0.00009		-0.00010	
SATV	-0.00006		-0.00001		-0.00007		0.00001	
HSGPA	0.738	***	0.754	***	0.759	***	0.768	***
HSRank	0.001		0.001		0.001		0.001	
FYRH	0.108	***	0.114	***	.1180	***	0.125	***
FG	-0.067	*			-0.007	*		
Core40					0.040		0.071	**
$20K-$40K	0.077				0.086			
$40K-$60K	0.038				0.032			
>$60K	0.092	*			0.084	*		
R^2	0.221		0.224		0.219		0.225	
adj R^2	0.218		0.222		0.215		0.223	

Table 5
Estimating the Financial Impact of All Programs

Annual Dollar Value of All Programs= $2,119,729	Total Number of Additional Students * Weighted Average of Tuition
Total Number of Additional Students = (164+134.4+164)=462.4	Number of Additional Sophomores+ Number of Additional Juniors + Number of Additional Seniors
Number of Additional Sophomores= (.08*2050)=164	Increase in First-Year Retention * Number of Freshman in a Typical Class
Number of Additional Juniors= (.819*164)=134.4	Conditional Second-Year Retention Rate * Number of Additional Sophomores
Number of Additional Seniors= (.847*134.4)=113.9 (.440*113.9)=50.1 113.9+50.1=164	Conditional Third-Year Retention Rate * Number of Additional Juniors + Conditional Fourth-Year Retention Rate * Number of Additional (Ungraduated) Seniors
Weighted Average of Annual Tuition= $4,584	(In-State Tuition * Percentage of First-Year students from Indiana) + (Out-of-State Tuition * Percentage of First-Year students from outside Indiana)

Table 5
Estimating the Financial Impact of Individual Programs

Dollar Value of Individual Program=	Dollar Value of All Programs *((Percentage Increase in First-Year Retention Attributable to the Program) / (Total Increase in Retention))
Learning Communities Low (.215*$2,119,729)= $508,735 Intermediate (.270*$2,119,729)= $561,728 High (.290*$2,119,729)=$614,721	
First-Year Residence Halls Low (.215*$2,119,729)= $455,742 Intermediate (.221*$2,119,729)= $468,460 High (.228*$2,119,729)=$483,298	
Percentage Increase in First-Year Retention Attributable to the Program	((Parameter estimate from Marginal Impact Table)*(Percentage of First-Year Population Participating in the Program))/ (Percentage Increase in First-Year Retention)
Learning Communities	
Low-end (.048*.4)/.08=.240	Lowest value of LC row in Table 3
Average (.053*.4)/.08=.265	Average value of LC row in Table 3
High-end (.058*.4)/.08=.290	Highest value of LC row in Table 3
First-Year Residence Halls	
Low-end (.033*.52)/.08=.215	Lowest value of FYRH row in Table 3
Average (.034*.52)/.08=.221	Average value of FYRH row in Table 3
High-end (.035*.52)/.08=.228	Highest value of FYRH row in Table 3