The Killer Course Hypothesis

(Working Paper)

Justin G. Gardner[1], Ph.D.
Middle Tennessee State University
School of Agribusiness and Agriscience
(615) 898-2418
jggardne@mtsu.edu

Selected Paper prepared for presentation at the Southern Agricultural Economics Association Annual Meeting, Corpus Christi, TX, February 5-8, 2011

[1] The author would like to thank the *Ad Hock Committee to Study the Impact of Gatekeeper Courses* at Middle Tennessee State University for their assistance with this project.

Abstract

Due to recent legislative changes, universities in Tennessee will receive funding based on student retention and graduation rates rather than enrollment. In light of these changes it is important that academics in all disciplines study retention rates in order to identify areas for improvement. I investigate the impact of "killer courses" on student retention both in the school of agriculture and in the general student population. In addition I explore alternative frameworks for addressing retention issues.

Introduction

Due to recent legislative action in Tennessee universities throughout the state will receive funding based on student retention and graduation rates rather than enrollment (THEC, 2010). In light of these changes it is important that academics in all disciplines study retention rates in order to identify areas for improvement. It is also important to educate policymakers on the dangers of incentivizing retention, persistence and graduation as opposed to incentivizing learning. As agricultural economists we are uniquely qualified to conduct research in order to inform policy makers about methods that can increase retention. In addition is it important to examine our own degree programs to ensure student learning. This paper is an effort to move in that direction. As such as I will review the literature on student retention, both inside and outside of agriculture and present a framework for conceptualizing the retention issue[2]. Finally, I will test the hypothesis that courses with low success rates (a.k.a. "killer courses") decrease the likelihood of retention.

Data to test this hypothesis will be drawn from registration records at Middle Tennessee State University (MTSU). MTSU is Tennessee's largest undergraduate institution with over 26,000 students. It is located in Murfreesboro, TN, the geographic center of the state. MTSU's school of Agribusiness and

[2] For the purposes of this paper, retention will be defined as returning to the same institution the following academic year, while persistence implies that the student will complete a degree.

Agriscience (ABAS) is home to approximately 450 undergraduate students majoring in Agribusiness, Plant and Soil Science, and Animal Science.

Literature Review

The literature on student retention is extensive, reaching back several decades (Tinto, 1975), and studies can be found in a wide variety of academic disciplines. Vincent Tinto, arguably one of the most influential researchers in this field, postulates a model of student retention that focuses on the student's academic and social integration. Tinto has identified four conditions that promote student retention (Tinto, 2002). The first condition is the university's commitment to retention, which entails an expectation that students will persist and eventually earn a degree. Students will rise to meet these expectations. Second, the university must provide academic and social support to the students. This may include things like providing a quality library (Emmons and Wilkinson, forthcoming), or developing a community within the classroom (Tinto, 1997, Tinto, 2002). Tinto's third condition is student involvement. Students that are academically and socially involved are more likely to stay and finish their degree. Tinto's fourth condition is learning. According to Tinto "The more students learn the more value they find in learning…."(2002, p3).

Cabrera et al. (1993) use a structural equations model to analyze student persistence. In their model they find significant pathways from encouragement from friends and family to Tinto's notion of academic and social integration. All of which lead to institutional commitment and goal commitment which in turn lead to a student's intent to persist. In their model grades both create and are created by academic integration and grades are influenced by financial attitudes. The intent to persist combined with grades leads to actual persistence. A structural equations approach is a significant improvement over single equation or single variable approaches as it implicitly models endogeneity.

In agriculture many retention studies can be found in the agricultural education literature. Cole and Fanno (1999) surveyed agricultural students that dropped out of college. Approximately fortyseven percent cited financial difficulty, while approximately forty-four percent changed their career goal. Thirty percent were not satisfied with the curriculum or the teachers and twenty-six percent cited poor progress toward graduation. It is apparent that students could select more than one reason; it is not apparent how these reasons interacted with each other.

Garton et al. (2001) used a step-wise regression procedure to analyze the importance of admissions test scores (ACT), high school core GPA, and rank in high school on college GPA. They also included results from the Group Embedded Figures Test (GEFT) which is used to assess student learning style; students could be field-neutral, field-dependent or field-independent learner. According to Garton et al. (2001) field-dependent learners prefer structure, are extrinsically motivated, think globally and have a difficult time solving problems. Field-independent learners are analytical problem solvers, prefer individual as opposed to group work, and are intrinsically motivated. They found that high school core GPA and ACT score are the best predictors of college performance. They also used a stepwise discriminate analysis to determine the best predictors of freshman retention. They found that the GEFT score and the ACT score were significant predictors, and field-independent learners are less likely to return for their sophomore year. Although the stepwise method is not valid for hypothesis testing the study points to some important results. First, much of a student's ability to succeed in college is a result of the academic resources they bring with them to college. Second, the authors assert that field-dependent learners perform better than field-independent learners in general education courses, adding that agriculture students are often exposed to more non-agriculture courses than agriculture courses during their first year on campus. Implying that tailoring the mix of courses to student needs can impact retention rates.

Koon et al. (2009) conducted a study using College Student Inventory (CSI) scores and a demographic survey. Rather than employing a regression approach they used Pearson correlation coefficients to find variables that are correlated with retention. The student's first year cumulative GPA was positively correlated with retention while family and emotional support was negatively correlated. The GPA result is expected, but the family support result is not, students that have more support should have a higher likelihood of being retained as Cabrera et al. (1993) showed. It is possible that this result is due to an unobserved component in the error term. Or rather it is due to an observable variable that should be included in a regression.

Leppel (2002), writing from the perspective of an economist, took a utility function approach to analyzing student retention and persistence. If the satisfaction, both present and future, from obtaining a degree exceeds the utility from dropping out the student will persist and finish the degree. As is usually the case with most temporal analysis Leppel (2002) assumes that future utility, resulting from a better job for example, is discounted. Students that are struggling, for any reason, are likely to quit school because the current utility drops. From here we can study known factors that influence persistence in a theoretical context. Students with lower family incomes are more likely to struggle because they have to work or take out student loans. Leppel (2002) hypothesizes that some factors could have positive or negative motivational effects, such as age and marital status. Students that are older or married may have higher motivational levels relative to traditional students, but an older student has a shorter time span to enjoy future benefits of a degree and a married student may have constraints that single students do not. Leppel (2002) found that age, marriage and hours of employment decreased the odds of persistence, while family income increased the odds of persistence.

Cameron (2009) proposed and tested the hypothesis that “killer courses” taken by first year students could have an adverse impact on student retention. Cameron defined “killer courses” as “courses which have a high proportion of students earning grades below C- and or withdrawing”(Cameron, 2009 p1). These courses may be viewed as gatekeeper courses. They are often rigorous courses which students must take as prerequisites for classes within their major or part of a general education curriculum. One might expect math and science to be “killer courses”, although

Cameron demonstrated that any number of classes could potentially be a “killer.” Cameron found that “killer courses” did not have an adverse impact on freshman retention, finding instead that students that took more classes of any type had a higher retention rate. As will be discussed below, this result may be driven by the way Cameron’s variable was constructed.

Two common threads seem to be present in the retention literature; the first is lax statistical standards. As an example consider Tinto’s analysis of the impacts of learning communities on retention. Tinto used a stepwise regression procedure which is inappropriate for hypothesis testing. The process inflates alpha levels thus increasing the chances of a Type-I error (Draper, et al., 1979). A Type-I error is a false positive, leading us to conclude that a relationship between a variable exists when it may not. Additionally, in the presence of collinear independent variables stepwise methods are unable to detect the best variables for inclusion (Studenmund, 2001). In the context of student retention we should expect collinear predictor variables, a student’s ACT score and high school GPA should be correlated. The resulting model may have no relation to any underlying theory. In disciplines where theory does not drive statistical modeling this is a non issue. But there is a significant body of retention literature that develops a theory of retention. Even though problems with the stepwise procedure should be well known this method still appears in published work such as Garton’s (2001) study outlined above. In the agricultural education literature one can find many published papers that do not employ enough sophistication to properly address the retention issues such as comparisons of conditional means (see Cole and Fanno, 1999, Cole and Thompson, 1999). It is common to find papers that demonstrate correlation between outcome variables and the factors believed to impact the outcome variables (see Koon, et al., 2009, Mezick, 2007). A correlation coefficient can only prove that two variables are correlated and the direction of that correlation. We cannot use a correlation coefficient to predict how a change in high school grade point average will impact retention. Leppel’s (2002) study however, is methodologically sound. She explicitly models endogeneity in her regression model. Other studies employ structural equation modeling which take endogeneity into account (Cabrera, et al., 1993).

The second common thread is the implicit assumption that there is a retention problem. Leppel (2002) remedies this by assuming that college students are utility maxamizers. The decision to return to school and persist through to the completion of a degree is the result of decision making by a rational individual, if the benefits outweigh the costs (including monetary cost, opportunity costs, and effort) the student will finish the degree. This is consistent with Spence's (1973) job market signaling game where individuals with a high marginal productivity can earn a degree with less effort, and thus are more likely to earn a degree. In Spence's classic model education must only serve as a signal. It can lead to higher wages even if it does not improve the individual's marginal productivity. These models imply that optimal retention and graduation rates may exist, with some individuals completing college and others not. In these models the results are driven primarily by attributes that the university cannot control such as reservation wages and student resource constraints.

A Framework for Analyzing the Retention Issue

The literature on retention implicitly assumes the existence of a retention problem. Increasing the number of college graduates is assumed to be an improvement for society. Better educated workers receive a higher wage and have a higher quality of life. Hence, higher retention rates must also be better. More students seeking degrees and higher retention rates lead to more tuition income, and the university is better able to grow and expand. From the perspective of universities that need student tuition income this may be the appropriate way to view retention. However, this view is inconsistent with a utility framework which assumes that students are rational decision makers that will only continue an education if the benefits exceed the costs. If we wish to increase retention rates the answer is simple, incentivize the student by lowering the costs and/or increasing the benefits. This, of course has many potential side effects that will be addressed below. A new framework is needed, research must seek to answer the appropriate questions and academics must educate policy makers on the role of higher education. It is important to make higher education accessible to many, but not everybody that has access to higher education should obtain a degree. I present two conceptual frameworks that can be used to break down the "retention problem" to real issues that can be addressed on an institutional level.

A Condorcet Jury Approach

Members of the university faculty can be thought of as a Condorcet jury. The Condorcet jury theorem states conditions in which majority rule is superior to a decision by a dictator. For an excellent exposition of the Condorcet jury theorem and some of its extensions see Dharmapala and McAdams (2003). The Condorcet jury can be viewed as a type of signaling game, where the signal is the evidence presented to the jury. It is commonly used in law and political science. Consider a student that has achieved the necessary learning objectives to pass a class, this achievement is evidence presented to the jury. If the evidence is high quality then as the size of the jury increases the probability that the jury will make the correct decision increases. The jury is increased by requiring students to take numerous courses with sufficient variety to ensure that evidence is presented to a large number of faculty members. In order for the Condorcet jury theorem to function the evidence must have a high (greater than 0.50) probability of being correct. It is generally assumed that the members of the jury receive the greatest utility when they make a correct decision (i.e. the jury convicts a guilty man or acquits an innocent man) and receives the lowest utility from making a mistake (i.e. the jury convicts an innocent man or acquits a guilty man). Thus, we can use the statistical concepts of Type-I error (false positive) and Type-II error (false negative) error to understand retention.

Assume that the null hypothesis is that a student has not achieved the learning outcomes needed to earn a college degree. If the null is true it would be a Type-I error to confer a degree. If the null is false it would be a Type-II error to deny the degree. This can be seen in Table 1. If earning a degree is a Spence-like job market signal then we can understand the dangers of committing a Type-I error. If employers cannot differentiate between high ability and low ability workers they will only pay low wages. Of equal importance, if a university develops a reputation for committing Type-I errors the reputation of the institution will decline and a degree from said university will be a signal of low quality. There are also consequences from committing an excessive number of Type-II errors. Increasing the number of Type-II errors does not imply that only high ability students will earn a degree, rather some high ability students will not earn a degree do not. Quality students that could earn a degree will learn of the university's' reputation and seek an education elsewhere, or choose not to go to college. The Tennessee state legislature implicitly believes that universities in Tennessee are committing an excessive number of Type-II errors as evidenced by the passing of the complete college act. It is my claim that increasing retention, while a noble cause, is misguided. It is better to focus on increasing learning in order to increase both Type-I and Type-II errors.

Table 1 Type-I and Type-II Errors

	The Student Graduates	**The Student does not Graduate**
H_0: The Student has *not* Achieved Learning Outcomes	Type-I Error	No Error
H_a: The student has Achieved Learning Outcomes	No Error	Type-II Error

In the context of a statistical test Type-I and Type-II errors are mathematically linked together. Holding the sample size constant as the probability of a Type-I error (α) increases then the probability of a Type-II error (β) decreases. As Tinto (2002) outlined one of the keys to increasing student retention is to increase learning. Thus, if we can remove barriers to learning we can increase graduation rates while decreasing both types of errors. If, as Cameron (2009) proposed, a university has an institution-wide learning barrier, such as "killer courses," in place then removing such a barrier would be prudent.

The danger with legislative action that ties funding to student progress is that the jury's payoff structure becomes altered; increasing the jury's utility from committing a Type-I error. With sufficient incentive the jury will confer a degree to all students. This will lower the reputation of the university and while it will increase the proportion of the population with a degree it will have no impact on the number of educated individuals. However, if such legislative action causes a university's utility of a Type-II error to decrease then the legislation may have the desired effect.

In the case of a university there are several interesting twists that can be explored. The jury sets the learning objectives and thus controls the quality of the evidence. Students can retake courses, thus giving them ability to repeatedly present evidence. Finally, students can select the jury by selecting individual courses or selecting a major.

A "Poverty" Model

Low achieving students may arrive on campus with learning barriers already in place, much the same way the poor are born into poverty. Poverty it is typically determined within the economic model, it is endogenous. One may model retention the same way that one might model poverty, explicitly accounting for endogeneity in statistical models. Like poverty, "academic poverty" is intertwined with a lack of resources, so we can use a model, similar to Ayoub (2008) to better understand retention.

Figure 1 The Vicious Circle of Academic Poverty

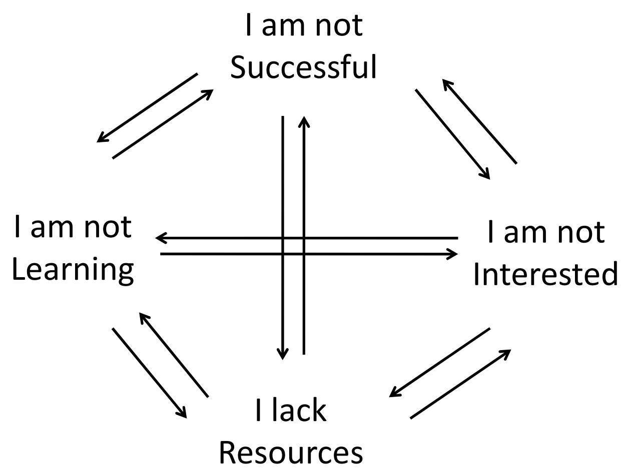

Figure 1 shows The Vicious Circle of Academic Poverty. According to Tinto (2002) students must learn in order to be successful and maintain interest, while a student that is not interested will not learn. This student will not commit the resources necessary to learn and success is not possible. A student that lacks goal commitment or institutional commitment (Cabrera, et al., 1993) will not be interested and thus will not earn a degree. A student that lacks resources will not earn a degree. I define resources broadly; they include prior academic training, on campus resources such as a quality library, an on-campus social network, support from family, and financial resources. The most direct way for a specific institution to address retention using this framework would be to increase tuition and increase admission standards thus filtering out students with low interest and limited resources. But, such a policy would be contrary to the goals of the Complete College Tennessee Act, which is to increase the proportion of the population with a college degree. Such a policy may be good for a single university, but not for an entire system.

An individual professor can do little to help a student with a resource deficiency, but a professor can stimulate interest and learning which will in turn incentivize the student to commit resources to learning. The university can also examine its own policies and procedures to look for barriers to learning. If, for example, the university were to create a series of courses that had no educational value beyond filtering out low quality students then it would run the risk of committing a Type-II error by decreasing student interest and hindering student learning. This is the essence of Cameron's (2009) killer course hypothesis which I propose to test.

Methods

Cameron (2009) proposed and tested the hypothesis that "killer" courses taken by first year students could have an adverse impact on student retention. Cameron defined "killer" courses as "courses which have a high proportion of students earning grades below C- and or withdrawing" (2009 p1). Often these courses are viewed as gatekeeper courses. These are rigorous courses which students must take as prerequisites for classes within their major or part of a general education curriculum. Typically, one would classify math and science as gatekeeper courses, although Cameron demonstrated that any number of classes could potentially be a "killer" or a gatekeeper course. My objective is to examine general education courses at MTSU in an attempt to determine if general education classes have an adverse impact on student retention.

The first and perhaps most difficult task is defining a "killer course." Cameron, when examining freshman retention, deemed any course with more than 150 students and a DFW rate greater than 10% a "killer course" (Cameron, 2009). This criteria appears to be to inclusive, at MTSU only 6 of the 59 general education courses have a D,F,W rate less than 10%. It may be reasonable to expect a D,F,W rate in excess of 10% for rigorous university classes, especially a university like MTSU that has a relatively open entrance requirements. This criterion also appears to be completely arbitrary as Cameron provided no justification for the 10% cutoff, and I am unaware of a method that can be used to construct a threshold that is not arbitrary. Moreover, I suspect that there exists some cutoff that would result in a statistically significant coefficient on a gatekeeper course variable.

Rather than mine the data to find an arbitrary cutoff that will "prove" that a given set of courses are gatekeepers I instead choose to test the hypothesis that general education courses serve as "killer courses" at MTSU. In other words I am trying to determine if students drop out as a result of encountering general education courses in the first semester of their freshman year. At MTSU general education comprises approximately one third of the hours required for degree completion. These courses have high enrollment, most have several hundred students in any given semester, and a handful have over 1,000.

Typically, retention studies have a binary dependent variable. If a student returns the following academic year then the retention variable is equal to one, and zero otherwise. Ordinary least squares

(OLS) regression is not appropriate in these situations. It is common to find logit models (Cameron, 2009), although probit models are also appropriate (Greene, 1997, Wooldridge, 2002). Emmons and Wilkinson (forthcoming) and Mezick (2007) use OLS to model retention at the intuitional level, using the institution's retention rate rather than an individual student's decision. OLS may be inappropriate in these situations because the dependent variable is bounded between zero and one, making a logistic regression a better tool. For my purposes either a logit or a probit model is appropriate. However, I would like to employ an instrumental variables method in my model. Since STATA contains a procedure for estimating instrumental variable probit models I choose to use a probit model. For an excellent overview of IV probit models see Ayoub (2008).

Data

Data were made available from Middle Tennessee State Universities' office of Institutional Effectiveness and augmented with United States Census Data (http://factfinder.census.gov). Variable definitions and variable means can be found in Table 2 and Table 3 respectively. LOAD is defined as the total number of hours taken by each full-time first-time freshman in the fall of 2008, while GELOAD is the number general education hours. GERATIO is the ratio of general education hours to total hours. ITGPA is the institutional term grade point average, i.e. it is the students MTSU GPA for fall of 2008. CLASSSIZE is the average number of students enrolled in the student's class. AA, HISPANIC, OTHER, FEMALE and PELL are indicator variables, set to 1 if the student is African American, Hispanic, some other minority, female or eligible for a Pell grant, respectively. HSGPA is the student's high school grade point average and COMP is the student's comprehensive ACT score. POP and MHI are the population of the student's home zip code and the median household income in the student's home zip code, based on the 2000 census.

Table 2 Variable Definitions

Variable	Definition
RETAINED	The proportion of fall 2008 first-time full-time freshmen that returned in the fall of 2009
LOAD	Number of hours taken in the fall of 2008
GELOAD	Number of general education hours taken in the fall of 2008
GERATIO	The ratio of general education hours to total hours.
ITGPA	Institutional Term GPA
CLASSSIZE	Average Class Size
AA	= 1 if the student is African American
HISPANIC	= 1 if the student is Hispanic
OTHER	= 1 if the student is some other minority
FEMALE	= 1 if the student is female
PELL	= 1 if the student is eligible for a Pell grant
HSGPA	High School GPA
COMP	Act composite score
POP	Population of the students home zip code (1999)
MHI	Median household income in the students home zip code (1999)

Table 3 Descriptive Statistics, 2008 First-Time Full-Time MTSU Freshmen

Variable	ABAS Freshmen			Non-ABAS Freshmen		
	Observations	Mean	Std. Dev.	Observations	Mean	Std. Dev.
RETAINED[$]	63	0.81	0.40	3221	0.73	0.45
LOAD*	63	14.79	1.46	3221	14.47	1.50
GELOAD***	63	9.16	3.14	3221	10.75	3.36
GERATIO***	63	0.62	0.21	3221	0.74	0.22
ITGPA	63	2.43	1.07	3218	2.57	1.08
CLASSSIZE	63	49.91	17.10	3221	50.04	21.89
AA	63	&	&	3221	0.19	0.39
HISPANIC	63	&	&	3221	0.02	0.15
OTHER	63	&	&	3221	0.04	0.20
FEMALE**	63	0.67	0.48	3221	0.52	0.50
PELL	63	0.25	0.44	3221	0.31	0.46
HSGPA	63	3.29	0.37	3221	3.26	0.51
COMP**	63	21.54	3.43	3099	22.52	3.53
POP***	63	18.49	14.84	3221	26.65	15.65
MHI	63	43.21	10.72	3221	45.88	14.87

[$]For reasons outlined in the text of this article, this is not the officially reported retention rate for MTSU
[&]Redacted due to the small proportion of minority students in ABAS.
*** p<0.01, ** p<0.05, * p<0.1

As the reader can see in Table 3, in the fall of 2008 there were 63 new full-time, first-time ABAS freshmen in this dataset, 81 percent of them returned in the Fall of 2009. Outside of ABAS, there were

3,221 new MTSU full-time first-time freshmen students included in this dataset, 73 percent were retained. The difference in retention rates between ABAS and non ABAS freshmen is not statistically significant. It is important to note that this is not the official retention rate for MTSU. Students who had completed a high school equivalency were excluded from the dataset because there is no straightforward way to compare a high school equivalency exam score to a high school GPA. Students under age 18 were not included in the dataset, and this dataset does not include first-time freshmen that did not take a full-time course load.

On average, all students took approximately 15 hours, roughly 10 of which were general education courses. During their first semester ABAS students spent roughly two-thirds of their course time in general education classes while non-ABAS students spent approximately three-fourths of their course time in general education classes. ABAS students tend to take more classes and fewer general education classes, the difference is statistically significant at the 10 percent and 1 percent level respectively. Minority students, including African Americans, Hispanics and Other minorities made up 25 percent of the general MTSU freshmen population. Few minority students are found in ABAS, results are withheld to prevent the possible disclosure of personal information. Approximately half of the freshmen in the general student population are female, while approximately two-thirds of the ABAS freshmen are female, a statistically significant difference. One fourth and 31 percent of the ABAS freshmen and non-ABAS freshmen were eligible for a Pell grant, respectively. ACT composite score, median household income and high school grade point average were similar for both ABAS and non-ABAS freshmen. ABAS freshmen tended to come from areas with less population, a difference that is statistically significant.

Results

Results for MTSU freshmen outside of ABAS are presented in Table 4. Model 1 is similar to Cameron's (2009) model in that it includes variables for both "killer courses," which in my model is the

GELOAD variable and all courses, LOAD in my model. Model 2 and 3 show how including both variables is inappropriate. If we drop LOAD from the model, as in model 3 we see that GELOAD becomes significant. If we drop GELOAD from the model, as in model 2 then the estimated coefficient for LOAD increases. This is because the two variables capture the same effect, as students take more classes they are more likely to be retained. To negate this effect I propose to use the ratio of GELOAD to LOAD in Model 4. If retention rates decline as the proportion of general education classes increases then we can say that these allegedly "killer courses" create a barrier to learning that negatively impacts retention. Model's 5 and 6 are instrumental variable models which attempt to take endogeneity into account. Student's self select the number of hours and the course mix. It is reasonable to believe that a student that lacks interest, or lacks resources will take fewer classes or reduce the proportion of classes thought to be "killer courses." All of the models in Table 4 are statistically significant, based on the χ^2 statistic. The pseudo-R^2 is about .23 for all models.

Table 4 Probit Results, Non-ABAS Full-Time First-Time Freshmen

	Non ABAS Full-time First-Time Freshmen					
VARIABLES	**1**	**2**	**3**	**4**	**5**	**6**
LOAD	0.0594***	0.0652***				0.285**
	(0.0196)	(0.0178)				(0.143)
GELOAD	0.00648		0.0181**			
	(0.00921)		(0.00832)			
GERATIO				0.156	0.679*	
				(0.130)	(0.400)	
CLASSSIZE	0.00170	0.00195	0.00141	0.00174	0.000751	0.00149
	(0.00129)	(0.00123)	(0.00128)	(0.00129)	(0.00168)	(0.00141)
ITGPA	0.678***	0.677***	0.688***	0.690***	0.692***	0.590***
	(0.0264)	(0.0264)	(0.0262)	(0.0262)	(0.0270)	(0.0902)
POP	0.00384**	0.00380**	0.00385**	0.00379**	0.00433**	0.00375**
	(0.00174)	(0.00174)	(0.00174)	(0.00173)	(0.00177)	(0.00176)
MHI	-0.000185	-0.000156	-0.000284	-0.000273	0.000365	0.000789
	(0.00183)	(0.00183)	(0.00182)	(0.00182)	(0.00190)	(0.00185)
PELL	0.0582	0.0563	0.0552	0.0514	0.0781	0.1000
	(0.0612)	(0.0611)	(0.0611)	(0.0610)	(0.0626)	(0.0625)
AA	0.383***	0.378***	0.398***	0.395***	0.412***	0.320***
	(0.0763)	(0.0760)	(0.0760)	(0.0761)	(0.0782)	(0.0908)
HISPANIC	0.0939	0.0913	0.115	0.116	0.239	0.139
	(0.183)	(0.183)	(0.183)	(0.184)	(0.194)	(0.195)
OTHER	0.257*	0.254*	0.260*	0.256*	0.300**	0.275*
	(0.143)	(0.142)	(0.142)	(0.142)	(0.150)	(0.149)
FEMALE	-0.0578	-0.0537	-0.0628	-0.0573	-0.0871	-0.0780
	(0.0543)	(0.0540)	(0.0542)	(0.0542)	(0.0572)	(0.0540)
CONSTANT	-2.199***	-2.226***	-1.475***	-1.418***	-1.798***	-5.229***
	(0.287)	(0.284)	(0.157)	(0.161)	(0.270)	(1.858)
OBSERVATIONS	3218	3218	3218	3218	3097	3097
PSUEDO R^2	0.236	0.235	0.233	0.232		
χ^2	886.3***	885.8***	877.1***	873.9***	709.5***	821.4***
χ^2 TEST FOR EXOGENEITY					2.613	2.049

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

In Model 1 and 2 the LOAD variable is statistically significant and positive, implying that students that take more classes have a higher likelihood of being retained. This is consistent with Cameron's (2009) findings. It is also consistent with Tinto's (2002) notion of academic integration; students that take more classes are more integrated and thus more likely to stay. It is also consistent with my notion of endogenous academic poverty. Students that lack interest take fewer courses and due to a lack of interest have lower chances for success. Likewise students that lack resources, rather they be prior academic training, or financial resources, take fewer classes. The lack of resources feeds into the viscous circle of academic poverty. In Model 3 the GELOAD coefficient is also statistically significant and positive. At first glance this model appears to refute the killer course hypothesis. *A priori*, one should suspect that this is because LOAD and GELOAD both are a function of the same underlying student behavior, students that are more likely to succeed will take more classes[3]. Hence it is not a useful model for testing the killer course hypothesis. I propose that using the GERATIO will be a better test. In Model 4 the coefficient on the GERATIO is not statistically significant, thus refuting the killer course hypothesis.

Models 5 and 6 employ an IV probit model to correct for endogeneity that may be present in the GERATIO and LOAD variables, respectively. The student's high school GPA and ACT composite score were used as instruments. The χ^2 test for endogeneity reject the null hypothesis that endogeneity is present in both cases. However, this could be due to poor instruments, high school GPA[4] and the ACT[5] composite score are both correlated with the student's first-term GPA (ITGPA). Thus, calling into question the assumption that the instruments are not correlated with the error term in the model.

Also of interest Table 4 shows that the coefficient on the population variable is positive and statistically significant, implying that MTSU freshmen from rural areas have a lower likelihood of returning to college for their sophomore year. These models indicate that African American students

[3] LOAD and GELOAD also happen to be correlated: r=0.41, p-value <.001
[4] r=0.47, p-value <.001
[5] r=0.34, p-value <.001

and other minorities have higher retention rates relative to white males in the general population. The student's grades, obviously, are statistically significant predictors of retention. However the estimated coefficient is likely to be endogenous. Since I am not concerned with testing any hypothesis regarding grades there is no need to correct for endogeneity in this case, ITGPA serves only as a control variable (see Gardner, et al., 2010, Heckman, 2001). One could imagine that students may find classes in large lecture halls to be discouraging, and thus such classes might have an impact on retention rates. However, CLASSSIZE was not statistically significant. Likewise variables to control for income (MHI and PELL) were not statistically significant, nor were dummy variables for Hispanic students and females.

For agricultural students the results were similar to the general student population. The models in Table 5 for ABAS freshmen are analogous to the models in Table 4 for non-ABAS freshmen. As can be seen in Table 5, population is not statistically significant. In Models 7 and 9 in Table 5 we see that GELOAD is positive and statistically significant, although only at the 10 percent level. LOAD is not statistically significant and in Model 10 the GERATIO variable is statistically significant, again at the 10% percent level. Implying that ABAS freshmen should increase the number of general education classes in order to improve their odds of returning. Of course, this result could be due to endogeneity. Models 11 and 12 correct for endogeneity, and like the general student population the χ^2 test for endogeneity fails to find that it exists, as in the general population high school GPA and ACT composite scores were the instruments. Again, these instruments may not be valid, ITGPA is correlated with HSGPA[6].

[6] r= 0.27, p-value = 0.03

Table 5 ABAS Full-time First-Time Freshmen

	ABAS Full-Time First-Time Freshmen					
VARIABLES	**7**	**8**	**9**	**10**	**11**	**12**
LOAD	-0.0752	-0.0214				
	(0.168)	(0.159)				
GELOAD	0.192*		0.185*			0.106
	(0.0995)		(0.0981)			(1.289)
GERATIO				2.383*	-2.403	
				(1.361)	(10.57)	
CLASSSIZE	-0.0127	-0.00076	-0.0133	-0.00993	0.0114	-0.00736
	(0.0161)	(0.0141)	(0.0160)	(0.0155)	(0.0527)	(0.0970)
ITGPA	0.689***	0.723***	0.695***	0.697***	0.823*	0.720**
	(0.241)	(0.224)	(0.241)	(0.237)	(0.426)	(0.309)
POP	-0.0128	-0.00577	-0.0118	-0.0118	0.00175	-0.00855
	(0.0182)	(0.0163)	(0.0181)	(0.0177)	(0.0343)	(0.0562)
MHI	-0.00181	-0.0123	-0.00089	-0.00294	-0.0165	-0.00397
	(0.0224)	(0.0213)	(0.0222)	(0.0219)	(0.0402)	(0.0527)
PELL	0.794	0.598	0.855	0.737	0.750	0.818
	(0.710)	(0.616)	(0.703)	(0.665)	(0.794)	(1.034)
BLACK						
HISPANIC						
OTHER						
FEMALE	0.456	0.393	0.427	0.456	0.277	0.410
	(0.473)	(0.451)	(0.468)	(0.465)	(0.621)	(0.594)
CONSTANT	-0.693	-0.0342	-1.764	-1.656	0.463	-1.302
	(2.871)	(2.731)	(1.569)	(1.569)	(4.843)	(7.811)
OBSERVATIONS	58	58	58	58	58	58
PSUEDO R^2	0.352	0.284	0.348	0.338		
χ^2	20.79***	16.77**	20.59***	19.98***	8.563	11.11
χ^2 TEST FOR EXOGENEITY					0.406	0.00397

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Due to the failure to detect and correct for endogeneity and the collinearly between LOAD and GELOAD the best models for predicting retention of first semester freshmen appear to be models 2, 9 and 10. These, along with marginal effects are presented in Table 6. For a Non-ABAS freshman, increasing the course load by one hour increases the probability or retention by 2 percent. For an ABAS freshman adding an additional general education course (i.e. a "killer course") increases the probability of retention by 3.5 percent. It should be noted that this result is only statistically significant at the 10 percent level. It is also important to note that African American freshmen at MTSU have a 10 percent higher chance of returning for their sophomore year, relative to white males.

Perhaps the most important result is the coefficient of the population variable in the Non-ABAS model. The result is important because the variable is not statistically significant for ABAS freshmen. The implication is that rural students have a lower retention probability when they do not choose to major in ABAS, while rural students that do are not handicapped. This could be a function of academic integration, perhaps rural students find it easier to connect with peers in the agriculture department.

First semester grade point average is a powerful predictor of student retention. In the general student population a 1 point increase leads to a 20 percent increase in the probability of retention. The effect is muted in ABAS where a 1 point increase leads to a 13 percent increase in the probability of retention.

Table 6 Marginal Effects

	Non ABAS		ABAS			
VARIABLES	**2**	**2 (MFX)**	**9**	**9(MFX)**	**10**	**10(MFX)**
LOAD	0.0652***	0.0200***				
	(0.0178)	(0.00546)				
GELOAD			0.185*	0.0352*		
			(0.0981)	(0.0184)		
GERATIO					2.383*	0.476*
					(1.361)	(0.278)
CLASSSIZE	0.00195	0.000599	-0.0133	-0.00252	-0.00993	-0.00199
	(0.00123)	(0.000379)	(0.0160)	(0.00310)	(0.0155)	(0.00314)
ITGPA	0.677***	0.208***	0.695***	0.132***	0.697***	0.139***
	(0.0264)	(0.00844)	(0.241)	(0.0495)	(0.237)	(0.0500)
POP	0.00380**	0.00117**	-0.0118	-0.00224	-0.0118	-0.00237
	(0.00174)	(0.000534)	(0.0181)	(0.00335)	(0.0177)	(0.00348)
MHI	-0.000156	-4.79e-05	-0.00089	-0.000169	-0.00294	-0.000588
	(0.00183)	(0.000561)	(0.0222)	(0.00420)	(0.0219)	(0.00435)
PELL	0.0563	0.0172	0.855	0.125	0.737	0.118
	(0.0611)	(0.0185)	(0.703)	(0.0762)	(0.665)	(0.0819)
BLACK	0.378***	0.106***				
	(0.0760)	(0.0191)				
HISPANIC	0.0913	0.0271				
	(0.183)	(0.0525)				
OTHER	0.254*	0.0711**				
	(0.142)	(0.0359)				
FEMALE	-0.0537	-0.0165	0.427	0.0870	0.456	0.0981
	(0.0540)	(0.0166)	(0.468)	(0.102)	(0.465)	(0.107)
CONSTANT	-2.226***		-1.764		-1.656	
	(0.284)		(1.569)		(1.569)	
OBSERVATIONS	3218	3218	58	58	58	58
PSUEDO R^2	0.235	0.235	0.348	0.348	0.338	0.338
CHI^2	885.8	885.8	20.59	20.59	19.98	19.98

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Summary

The evidence presented in this study supports Cameron's (2009) findings that "killer courses" do not impact retention rates. Of equal importance, this study examines the question with a great deal more statistical rigor, something that is lacking in the student retention literature. In addition I propose a new framework for analyzing retention. First, when analyzing retention it may be fruitful to treat low academic performance in the same manner one would treat poverty. Second, I propose that student learning is more important than retention. Retention studies should focus on identifying barriers to learning, such as "killer courses." In doing so researchers can focus on Type-I errors, granting a degree to students that have not achieved learning objectives, and Type-II errors, failing to grant a degree to students that have achieved learning objectives. By removing learning barriers retention rates can increase while simultaneously reducing the probability of both types of errors.

References

Ayoub, A. S. "Modeling Natural Resource Scarcity and Poverty Effects On fertility in Honduras, Nepal and Tanzania." University of Nevada, Las Vegas, 2008.

Cabrera, A. F., A. Nora, and M. B. Castaneda. "Structural Equations Modeling Test of an Integrated Model of Student Retention." *The Journal of Higher Education* 64, no. 2(1993): 123-139.

Cameron, J. "The Attack of the Killer Courses." Consortium for Student Retention Data Exchange Annual Symposium.

Cole, L., and W. Fanno. "Survey of Early Leavers: Implications for Recruitment and Retention." *NACTA Journal* 43, no. 1(1999): 53-56.

Cole, L., and G. W. Thompson. "Survey of Current Students: Implications for Recruitment and Retention." *NACTA Journal* 43, no. 3(1999): 15-20.

Dharmapala, D., and R. H. McAdams. "The Condorcet Jury Theorem and the Expressive Function of Law: A Theory of Informative Law." *American Law and Economics Review* 5, no. 1(2003): 1-31.

Draper, N. B., I. Guttman, and L. Lapczak. "Actual Rejection Levels in a Certain Stepwise Test." *Communications in Statistics - Theory and Methods* 8, no. 2(1979): 99-105.

Emmons, M., and F. C. Wilkinson. "The Academic Library Impact on Student Persistence." *College and Research Libraries* (forthcoming).

Gardner, J. G., R. F. Nehring, and C. H. Nelson. "Genetically Modified Crops and Household Labor Savings in US Crop Production." *AgBioForum forthcoming*(2010).

Garton, B. L., J. E. Dyer, and B. O. King. "Factors Associated with the Academic Performance and Retention of College Agriculture Students." *NACTA Journal* 45, no. 1(2001): 21-27.

Greene, W. H. *Econometric Analysis*. 3 ed. Upper Saddle River, New Jersey: Prentice-hall, 1997.

Heckman, J. J. "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture." *Journal of Political Economy* 109, no. 4(2001): 673.

Koon, L. A. F., M. J. Frick, and C. G. Igo. "What Kind of Students are Enrolling in a College of Agriculture and are they Staying?: A Mixed Methods Approach." *NACTA Journal* 53, no. 2(2009): 21-28.

Leppel, K. "Similarities and Differences in the College Persistence of Men and Women." *The Review of Higher Education* 25, no. 4(2002): 443-450.

Mezick, E. M. "Return on Investment: Libraries and Student Retention." *The Journal of Academic Librarianship* 33, no. 5(2007): 561-566.

Spence, M. "Job Market Signaling." *The Quarterly Journal of Economics* 87, no. 3(1973): 355-374.

Studenmund, A. H. *Using Econometrics: A Practical Guide*. 4 ed: Addison Wesley Longman, Inc., 2001.

THEC (2010) "Complete College Tennessee Act of 2010", http://www.state.tn.us/thec/complete_college_tn/complete_college_tn.html, Accessed on September 15, 2010.

Tinto, V. "Classrooms as Communities: Exploring the Educational Character of Student Persistence." *The Journal of Higher Education* 68, no. 6(1997): 599-623.

Tinto, V. "Dropout from Higher Education: A Synthesis of Recent Research." *Review of Educational Research* 45, no. 1(1975): 89-125.

Tinto, V. (2002) Promoting Student Retention: Lessons Learned from the United States. Prato, Italy.

Wooldridge, J. M. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: The MIT Press, 2002.