Some Remarks on Astin's Persistence Model

William Rosenthal
MI/AIR Meeting
October, 1994

Introduction

Retention has become something of an industry within the education business. At this date there is little point in reviewing the extensive literature in the field, but there are three issues which dominate that literature: retention as a method for maintaining enrollments in times of smaller high school graduating classes, retention as desirable for the individual student, retention information required by the Right to Know act and expected as part of the North Central Associations assessment mandate.

Although the first two constitute two of the sacred beliefs of our time, both seem to me to have potential flaws. Surely, an institution of higher education which managed to drive off most of its students would be suspect, but what about the institution which retains most of its students in the face of no evidence of achievement? Parsons College, late of Fairfield, Iowa comes to mind as the extreme example. It maintained a high-rolling lifestyle (including high faculty and administrator salaries) through low admissions standards and an amazing retention rate. The ride ended when Parsons lost its accreditation. On the other side of the coin was the Horticulture program at St. Louis Community College at Meramec which, at one time, had a retention rate so low as to merit serious concern on the part of SLCC planners. During a program review process ,assessment minded administrators discovered that virtually all students in the program were offered jobs at the end of the of the first year of the two year program. The program was highly rated by the local floral, landscaping, and lawn care industries and by its students but had a terrible looking persistence rate. The students, in accepting offers consistent with their goals for enrolling in the program, made it clear that they believed that retention in the program was not in their best interests. These examples are extreme; however, I hope that they provide the grain of salt which should be taken with all discussions of retention, particularly those which attempt to compare rates.

The Alexander Astin article which is the focus of this paper makes some assumptions about retention which should be questioned before one draws conclusions based on its methods. Its title "How good is your Retention Rate?" assigns a value to retention which ought to be questioned and discussed before numbers produced from the Astin model (or any other model) are interpreted. In their 1980 book on retention, Lenning, Beal, and

Sauer are very clear:

> To assume that retention equates with success and attrition with failure poses hazards for any retention study... But for any attrition or retention study to be worthwhile, this dichotomous notion should be replaced with a more objective understanding of what enrollment, graduation, and other kinds of statistics indicate, along with a commitment to develop programs that can help students reach the best decision about leaving or staying in school

These remarks aren't meant to suggest that a retention yardstick is a bad idea or that Astin's point (see attached essay and article) that persistence rates have no meaning outside a context is a bad one. Quite the contrary, it is a good one which may be helpful in thinking about some elements of student outcomes. (One might also note that it came about in response to the "Student Right to Know.." act, which assumed that rates had meaning in and of themselves.) My concern is that the "yardstick" of inputs as suggested by Astin is but one element of context. It is a yardstick guaranteed at the cost of a little arithmetic to provide an answer to an important question of measurement related to retention (should we have had higher numbers given the material we started out with). Not to demean the question, but one should remember Parsons College, just to keep a sense of perspective. Think about how good their retention rates would have looked next to their students "input" records. And what shall we do about accounting for different rates program by program on a campus? In other words, persistence, in and of itself, is not an adequate outcomes measure. It must be viewed in the light of other goals and issues. To make a persistence yardstick meaningful, other questions must be asked. This paper will travel only a short distance down that path, but it attempts to address the issue of some possible next steps.

In a 1993 paper (attached) Astin proposed a yardstick for judging an institution's undergraduate persistence rates against a calculated standard for a given set of incoming freshman characteristics. (In effect, he asks the analysis of covariance question, is this rate meaningful over and above the characteristics the student brought with her/him.) His approach uses data from 39,243 students at 129 colleges and uses regression statistics to model persistence behavior. There is potential for considerable debate on the amount of variance not accounted for in his models; however, in his 1976 paper (attached), Haase discusses a Bayesian model for prediction similar in its assumptions to Astin's. Haase also predicts from that model on the basis of a very small R^2.

Rather than haggle over fairly esoteric assumptions, a pragmatic approach is to try the model on local data and to regard the results as material for discussion as opposed to reasons to start writing press releases. In the process of developing and discussing local calculations of Astin's model, our discussions moved from an inability to define what the comparison meant to the question of whether we might not be able to use Haase's rather austere Bayesian model and data from our own recent past to develop predictive models of our own. A local model has the demerit of not having the compelling ability to allow an institution to claim superiority over other institutions. On the other hand, viewing persistence patterns in a familiar situation seemed promising and might permit an attempt at adding another important element of context, standards. "Standards" as I mean the term includes "how tough they grade" but is not limited to that. It also includes the difficulty of the curriculum. This is the next phase of the project and is not a part of this paper; however, it is easy to see that quality comparisons become possible along with comparisons of entering characteristics as one models the characteristics and behaviors of students in programs for judgements about program quality and difficulty could be made. For example, comparing persistence and context issues for an engineering college of good reputation with other colleges on its campus would allow one to ask not only the Astin question (given the entering characteristics of the students how do the persistence rates compare), but also questions about how the rates compare given the ability of the programs to attract students with characteristics which match the demands of the program, difficulty of the curricula, the success of the programs' graduates in moving on to the next stages of their careers, and so on. Many of the issues which should be a part of an assessment activity become part of the process of developing the local model.

To present both the original intention of this project and the first stages of the project which is developing from it, I shall present and discuss tables which compare institutional persistence rates to projections based on Astin's study. Please remember that no claims are being made as to "goodness" or "badness". The second part of the paper discuss models using local data and Haase's statistical approach.

One tempting question relates to persistence for the College Achievement Admissions (CAAP) program, the special admissions program for minority and disadvantaged students who show promise but do not meet MSU's regular admissions criteria. There has not been a good method for judging the persistence rates for students who would not have been admitted outside the CAAP program. One school of thought has it that their projected persistence rate should be zero, since they would not have been admitted, but, in fact, had these students not shown promise, they would not have been admitted under the CAAP rules. To compare any measure of success for this group to those for the regularly admitted group makes no more sense than matching up an intramural flag football team against the varsity team. In both cases, we would ask how well they did, given that they started out at a disadvantage. Answers to that kind of question are, of necessity, pretty subjective. The Astin model appears on its face to provide a less subjective method of making such a judgement, but its data base is so broad that these comparisons may be of the club versus the varsity variety. If, in fact, its results are useful, the implications for assessment are clear: overachievement should trigger a look at whether the program is doing something very, very right or if it is getting students through without a meaningful education; underachievment should trigger discussion about how realistic the retention goals are, need to improve retention, how well the students are being advised, how well they are being instructed and supported, and many others.

Method

I have selected three year's entering first time freshmen (fall terms, 1986, 1987, and 1988) on the bases that enough years have passed to have four year rates and that not so many years have passed that major changes in conditions render the results inapplicable. The cohorts are combined to avoid minor annual fluctuations. The combined population is divided into those regularly admitted and those admitted under the rules for the CAAP program. The Astin model also requires identification of black and Native American categories. Because the Native American population at MSU is small, I have not run separate models for them; however, I have run models for black students in part because previous studies at MSU have shown there to be quite different persistence rates for regularly admitted black students and CAAP black students. I chose the four year degree or completion of the fourth year model (Table 2, Astin) in that MSU's pattern is one of roughly half its bachelors' degrees being granted in the fourth year, the other half in the fifth year.

In addition to running the models using Astin's conventions, I have also used Haases's Bayesian model (article attached). One of the interesting features of Haase's model is that the results are presented in a percent based probability scale, so that one may predict that a student of given entering characteristics will have a, for example, 53 percent of finishing the fourth year in program "A" but a 68 percent chance in program "B". This ability to predict success may prove valuable in helping students to find programs in which they can persist. These models were calculated using high school gpa and a modified high school gpa as used at MSU for projecting first semester college gpa. This method modifies high school gpa according to a formula which takes into account actual MSU gpas for students from a given high school and gives as the result a projected MSU gpa which is a percentage of the high school average. In the process of deriving the projected gpa these formulas produce a coefficient for each high school which can be used to adjust high school gpa to reflect the performance typical of students from that high school. The modified gpa does not work well in the Astin model because it is being compared to the unmodified gpa's gathered in the Astin study and simply results in lowering the predictions.

Results and Discussion

Table 1 shows projections from Astin's table 2 full model and GPA only model (bachelor's degree or four year's enrollment) for regularly admitted students and the actual fourth year retention rates. Results of a model using Astin's formula and MSU data are at the far right of the table.

Projections from an Astin model accounting for GPA, SAT scores (ACT scores are converted to SAT via an Astin conversion table), and gender and an Astin model accounting only for GPA are presented.

MSU's retention rates are higher than the Astin projections. Whether that is good or bad is a matter for discussion and investigation of the kind described above. One possible conclusion, if one were to take the Astin sample as an appropriate comparison group, is that major increases in retention rates which are already higher than predicted will require more than ordinary measures.

Table 2 displays results from the Haase calculations. For regularly admitted students, both the adjusted and unadjusted GPA models do a fairly accurate job of modeling persistence for regularly admitted students. The model using adjusted GPA loads considerably more of its prediction on GPA than does the unadjusted model suggesting that the adjusting the GPA takes into account otherwise unknown or difficult to measure characteristics of student populations. I have include at the far right as set of figures from running the Astin regression model on the same data. It models overall persistence slightly lower than the other two models, which would, again, tend toward comparisons more favorable than warranted. It loads less of its prediction on the GPA than the Haase model using unadjusted GPA.

Where does this leave us? With a lot more work to do:

The next steps should include the following:

1. running predictive models at many combinations of entering characteristics by college or, perhaps, major.

2. comparing results according to beliefs about program difficultly as well as entering characteristics.

3. using results to project and evaluate performance for freshman

classes which entered after the years on which the predictive models are based.

Table 1

Degree or still enrolled at end of 4th year
Unadjusted MSU GPA

Group	Astin Projection[1] "full" model	Astin Projection[1] GPA only	MSU Actual[2]	Astin Model MSU DATA[3]
CAAP	42.46	51.48	53.06	44-46
Reg Admit Black	54.11	57.31	60.52	54
Reg Admit White	61.67	62.03	74.67	70
Reg Admit All	64.25	60.22	72.56	70

---

[1] From Astin's table 2

[2] Source: Persistence Rates for Domestic Undegraduate Students, 1994, table 3.2. Note that these rates are averaged results for multiple entering cohorts.

3. Derived figures for whole group from male/female model.

Table 2
Bayesian Prediction Model
Comparing unadjusted and adjusted GPA
Regular Admits

	unadjusted gpa	adjusted gpa	Astin Model MSU DATA	
constant	64.71	54.75	68.94	
gpa @2.0	7.45	18.34	3.89	
SatMath @470	3.27	2.67	—	
SatVerbal @420	.60 (75)	.48 (75.98)	—	72.83
female	-2.87 (72.4)	-3.21 (72.75)	- 4.2	68.63
black	-13.41(59)	-11.08 (61.78)	- 12.0	56.63

Point of View

By Alexander W. Astin

College Retention Rates Are Often Misleading

CONGRESS passed the Student Right to Know and Campus Security Act in 1991 as part of a growing legislative interest in making colleges and universities more "accountable." Although the act requires institutions to make certain disclosures regarding student athletes, campus-security policies, and crime statistics, perhaps its most provocative directive requires an institution to "disclose . . . its completion or graduation rate of full-time certificate-seeking or degree-seeking undergraduate students." Institutions were required to begin releasing these data in July, so it is reasonable to ask now what the consequences are likely to be for institutions and students.

Unfortunately, studies recently completed by the Higher Education Research Institute at the University of California at Los Angeles suggest that this reporting requirement is ill conceived and could have negative consequences for both institutions and students.

The proposed rules for implementing the statute, published in the *Federal Register* on July 10, 1992, imply that colleges and universities can be made more "accountable" by providing "consumers" (students and parents) with information to help them choose among postsecondary institutions: "These proposed regulations would require an institution to make completion and graduation rates . . . available to current and prospective students. . . . The Secretary also encourages institutions to make the completion and graduation rates available to secondary schools and guidance counselors so they have the information needed to advise student and parent consumers."

This language clearly implies that the data will be useful in helping students make decisions about where to attend college. But how can such information help? The not-so-subtle implication of the law is that the higher the retention rate, the better. Or, to put it more bluntly: Institutions with high retention rates are presumably doing a "better" job of retaining their students than are institutions with lower rates. In short, the prospective student is being encouraged to avoid institutions with low rates and to prefer institutions with high rates.

The recent research done at the institute suggests, however, that a simple retention "rate" tells us a lot more about who an institution *admits* than about how effective its retention practices are. Our longitudinal study involves data from 39,243 students attending 129 four-year colleges and universities. Regardless of where they attend college, the least-well-prepared students (those with C averages in high school and SAT composite scores below 700) are *five* times more likely to drop out (86 per cent versus 17 per cent) than are the best-prepared students (those with A averages and SAT scores above 1,300). Thus, institutions that admit large numbers of less-well-prepared students will tend to have low retention rates, and those with well-prepared students will tend to have high rates, *regardless of how effective their retention programs are.*

Formulas derived from multiple regression analyses using our entire sample of 39,243 students show that high-school grades and SAT scores carry the most weight in predicting who will complete college, but that other characteristics of entering students, such as race and sex, also carry some weight. For each of our 129 institutions, we used these formulas to compute an "expected" retention rate based upon the high-school grades, admissions-test scores, sex, and race of each entering student. By comparing this expected rate with the actual rate, we get a much better indication of how "effective" an institution actually is in retaining and graduating its students.

The most effective institutions will have actual rates that substantially exceed their expected rates, while the least effective ones will show the reverse pattern. Institutions with average retention will have similar expected and actual rates. Since *more than half* of the variation in retention rates among the 129 institutions can be explained by their expected retention rates, most selective institutions will have "good" retention rates even if their retention programs are mediocre, and most non-selective institutions' rates will be "poor" unless they happen to have exceptionally effective retention programs.

The danger in looking only at simple retention rates can be illustrated with an example from our 129 institutions. Let's look at two institutions whose actual retention rates are very different: A private university that graduated 54 per cent of the students it admitted and a historically black college that graduated 36 per cent. Taking these data at face value, a student, parent, or counselor would conclude that a student's odds of graduating are better at the private university. However, if we also look at expected rates, we reach a very different conclusion.

Whereas the expected rate at the private university is

ROBERT SOULE FOR THE CHRONICLE

67 per cent, the expected rate at the black college is only 22 per cent. So the private university *reduces* its students' chances of retention by 13 percentage points (67 minus 54), whereas the black college *increases* its students' chances of completing college by 14 percentage points (36 minus 22).

Clearly, looking only at actual retention rates, as the statute encourages, provides a misleading picture of how attending these institutions actually *influences* the student's chances of graduating. Not only is the private university's higher retention rate entirely attributable to its more selective admissions policies, but an individual student's chances of graduating are actually better at the black college. I could cite dozens of other examples of misleading conclusions that might be drawn from actual retention rates.

WHEN the right-to-know law was being debated in Congress, some higher-education officials argued that it would be "unfair" to compare retention rates of different types of institutions. What our research does is to circumvent this problem by developing an *internal* standard (the expected retention rate) against which the institution can judge its own performance (the actual rate). In effect, the institution is being compared with itself: "How well are we doing, given the students we admit?"

One of the most insidious features of the Student Right to Know Act is that, by focusing entirely on a student "outcome" measure such as graduation and by ignoring the student's original "input," the law creates a strong *dis*incentive for institutions to recruit underprepared students, not to mention poor students and those from underrepresented minority groups (who tend to be less well-prepared academically). If an institution is interested simply in maximizing its graduation rate (a strategy that the law implicitly encourages), that institution would have little to gain and potentially much to lose by admitting significant numbers of underprepared students. The law is, in effect, discouraging institutions from recruiting and admitting poor and minority students.

Our research suggests that student "consumers" should resist basing their college choice on raw retention rates and should instead seek to find out an institution's actual and expected retention rates. Although our institute has recently prepared a set of tables (which I can provide upon written request) that will allow any baccalaureate-granting institution to compute its expected retention rate, still other difficulties exist with the Student Right to Know Act that are not so easily resolved. For example, even when we compare actual with expected retention rates, the relative "effectiveness" of different institutions can be substantially affected by the particular definition of retention used.

FOR EXAMPLE, one major public university in our sample is clearly doing a poor job of graduating its students in four years: Its expected rate of 64 per cent compares with an actual rate of only 35 per cent. However, when we also count as "retained" those students who are still enrolled and those who completed four years of study without graduating, the expected and actual rates are exactly the same, 79 per cent. Thus, the six-year retention rates required by the statute will penalize some institutions and favor others, even if we are able to compare actual with expected rates.

Retention is not the only "outcome" measure that can lead to erroneous conclusions about institutions' effectiveness. Other examples would include the "pass rate" on tests such as the National Teacher Examinations, the state bar exams, and Florida's "rising junior" exam, or even the default rate on federal student loans.

It is unfortunate that the Student Right to Know Act was passed in its current form, but we need to realize that educational researchers and measurement specialists have been encouraging policy makers to use just such "outcomes only" assessments for many years.

Practically every school district in the country uses standardized tests to provide output measures, and the public is encouraged to believe that schools generating the highest (outcome) scores on statewide or district-wide tests are the "best," while those with the lowest scores are the "worst." Such quality judgments are meaningless without "input" data on the students when they first enroll. In fact, "outcome" scores are probably telling us much more about the population *recruited* by the school than they are about the effectiveness of the school's academic program. The same mindless form of one-shot testing also characterizes the National Assessment of Educational Progress and the testing now being proposed by the National Education Goals Panel.

It is a shame that educational researchers continue to support (and profit from) such simplistic and naive assessment practices, and that we continue to encourage policy makers and "consumers" to use the questionable results generated by them. The assessment community must produce and disseminate to policy makers and parents much better tools for evaluating the quality of our colleges and universities.

Alexander W. Astin is professor of higher education and director of the Higher Education Research Institute at the University of California at Los Angeles.

How Good is Your Institution's Retention Rate?[1]

Alexander W. Astin
University California, Los Angeles

In a recent paper criticizing sections of the federal "Student Right-To-Know" act (Astin, 1993a), it was shown that an institution's undergraduate retention rate can be a very misleading indicator of its capacity to retain students. Indeed, more than half of the variance in institutional retention rates can be attributed directly to differences in the kinds of students who initially enroll, rather than to any differential institutional "effect." The study also showed that some institutions with "high" retention rates should really have rates that are even higher, given the kinds of students they admit. By the same token, a number of other institutions with retention rates that appear to be very modest are actually retaining their students at a significantly higher rate than would be expected from their student inputs.

The purpose of this note is to provide individual institutions with the capacity to evaluate their own retention rates. The procedure basically allows any individual institution to calculate an "expected" retention rate based on the characteristics of its entering students. If the expected and actual retention rates are close, it can reasonably be concluded that the institution's capacity to retain students is on a par with other institutions nationwide. Institutions that are unusually adept at retaining students would be expected to have actual retention rates that substantially exceed their expected rates, whereas those with weak retention capacity would have retention rates that fall below the expected rate.

Calculating An Expected Retention Rate

The data used for calculating expected retention rates are derived from the Cooperative Institutional Research Program's 1985 entering freshman survey. Four and a half years after the students entered (during the 1989-90 academic year), selected institutions participating in the

---

[1] The author is indebted to Linda Sax and William Korn for their assistance in many phases of this project.

survey were asked to provide retention information on randomly selected samples of their 1985 freshman. Data were eventually obtained from a total of 39,243 students from 129 four-year colleges and universities, an average of 304 students per institution. For each of these students, three different dichotomous retention measures were constructed:

- A "retained" student is one who had earned the bachelor's degree at the time of the 1989-90 follow-up (score 1); non-retained students include all others (score 0).
- A "retained" student is one who had earned a bachelor's degree or had completed 4 years of undergraduate work at the time of 1989-90 follow-up (score 1); non-retained includes all others (score 0).
- A "retained" student is one who had either earned the bachelor's degree, completed 4 years of undergraduate work, or was still enrolled at the time of the 1989-90 follow-up (score 1); non-retained students include all others (score 0).

The formulas for deriving an expected retention rate for an institution were developed through a series of multiple regression analyses in which one of the retention measures (scored as 1 or 0) served as a dichotomous dependent variable and the student's high school grades, admissions test scores, sex, and race were used as independent (predictor) variables. Although a number of other entering freshman characteristics add significantly to the prediction of retention (Astin, 1993b), these four variables account for the bulk of the variance in retention that can be predicted from entering freshmen characteristics. The formulas reported here will be limited to these four input variables since it is likely that most institutions have information about these variables on their entering students. Information about other input variables that add to the prediction of retention can be obtained from the original study (Astin, 1993b). Other input variables that contribute independently to the prediction of retention include socioeconomic status, religion, hedonism, and political orientation (see Astin, 1993b, pp 193-194).

Table 1 shows the formulas for predicting the most stringent retention measure—completing a bachelor's degree within four years after entering college—using four different sets of input characteristics. Formula 1 is the simplest, employing only the student's average grade in high school. Note that <u>high school grades must first be converted to the same coding scheme shown in footnote "b" of Table 1</u>. (This grade conversion must be done regardless of which formula is used.) Thus, to estimate the student's chances of completing a degree in four years using only high school grades, the first formula would be applied as follows:

Probability of completing a degree in four years	=	a	+	b (high school grades)
"	=	.0069	+	.0915 (high school grades)

For example, if a student has an average grade of A- (code = 7; see footnote "b" in Table 1), you would multiply 7 by .0915 and add the result .0069, yielding a probability of .647. In other words, roughly two-thirds of the students who enter college with a high school grade average of A- complete college within four years after entering. By contrast, if the student's average grade in high school is C- (code = 2), the probability of completing a degree in four years is 2 x .0915 + .0069, or .197. Thus, an entering freshman with an average high school grade of C- has only about one chance in five of finishing college within four years.

Formula 2 in Table 1 is for use by institutions that have <u>both</u> high school grades and college admissions test scores available for their students. Institutions that use the ACT rather than the SAT can use the conversion table shown in the attached appendix to convert the ACT subtest scores into equivalent scores on the SAT verbal and SAT math tests. The use of formula 2 follows once again the usual regression formula, except in this case there are three predictor variables, each with its own coefficient. Formula number 2 thus looks like this:

Probability of completing a degree in four years	=	a	+	b_1 (grades) + b_2 (SAT-M) + b_3 (SAT-V)
"	=	-.2729	+	.0622 (grades) + .000454 (SATM) + .000433 (SATV)

Let's say we have an outstanding freshman with an A- average from high school (code = 7) and SAT scores of 750 and 650, respectively, on the math and verbal tests. Multiplying each of these three input variables by its respective coefficient, summing the products, and adding the constant yields a probability of .7845. Thus, a little better than three-fourths of freshman who enter college with such academic credentials would be expected to earn a bachelor's degree within four years. On the other hand, applying the same formula to a freshman who enters college with only a C average (code = 2) and SAT scores on the math and verbal tests of only 400 and 450, respectively, yields a probability of only .2280. In other words, slightly less than one student in four who enters college with such grades and test scores would be expected to complete college within four years. While the multiple correlation involving these three variables is only .327 (accounting for slightly more than ten percent of the variance in retention), these two hypothetical students have very different chances of completing a degree within four years. Thus, the student with high grades and test scores is three times more likely to complete college (78 percent) than is the student with low test scores and grades (23 percent).

Similar procedures should be followed in using formulas 3 and 4. Formula 3 is available for those institutions that have gender data on their students, whereas formula 4 is available for those institutions that have gender as well as racial/ethnic data. An important point to remember about using gender and race data is how these variables are coded: these "dummy" variables are coded either "2 or 1," rather than the traditional 1 and 0 (see footnotes "d" and "e" in Table 1). Special attention should be paid to the racial variables, since it is essential that each student receive a score on all four race variables. In other words, a white student would receive a score of "2" on the variable Race: white and scores of "1" on the three other race variables. A student who is from some racial group other than the four shown in Table 1 should receive a score of "1" on all four race variables.

Note that the multiple correlation coefficients shown for each of the four formulas in Table 1 increase slightly with the addition of more variables (from .288 in formula 1 to .334 in formula 4). What this means is that the accuracy of the prediction is increasing slightly as additional variables are added to the equation. Although the racial variables increase the multiple correlation by only a very slight amount (.004), note that the b coefficients for the four race variables suggests that race can make a substantial difference in the student's chances of finishing college in four years. Among students of the same sex and with identical high school grades and test scores, a white student, for example, would have a .146 better chance of finishing college in four years than would a Native-American student (.0250 + .1214), and an African-American student would have a .0654 better chance than would a Native-American student (.1214 -.0560).

Investigators wishing to use less stringent retention measures should employ the formulas shown in Tables 2 and 3. Note, however, that the multiple correlation coefficients (R) decline as the definition of retention becomes increasingly liberal. What this means, in essence, is that the most stringent measure—completing a bachelor's degree within four years—is easier to predict than the other two more liberal measures. This finding is consistent with a national retention study done nearly twenty years earlier (Astin, 1975), which showed that students who take longer than four years to complete a bachelor's degree more closely resemble the permanent drop-outs than they do those who complete the degree within four years. In other words, while the more stringent measure will misclassify as "non-persisters" many students who will eventually complete their degrees, the less stringent measures will incorrectly classify as "persisters" many students who will become permanent drop-outs. Readers should also keep in mind that there is simply no way to obtain a "perfect" measure of retention until all students have either completed their degree or died.

Computing an Estimated Retention Rate

Investigators desiring to compute an estimated retention rate for any entering cohort of students are advised to follow a four step procedure:

1. Decide which retention measure is most appropriate for your purposes (Table 1, Table 2, or Table 3).
2. Choose the formula (1, 2, 3, or 4) that suits the data that are currently available on the entering cohort.

3. Using the appropriate formula, compute for each student in the cohort an estimated probability of retention.
4. Calculate an expected retention rate for the entire cohort by averaging the individual probabilities.

If mean scores for the cohort are available on the relevant input variables, the computational process can be greatly simplified, since multiple linear regression is an "additive" model. Thus, all one needs to do is to multiply each mean by its respective coefficient, sum the products, and add the "a" constant. In taking this short cut it is important to realize that means for race and gender will range between 1.0 and 2.0. Thus, if the entering cohort includes 60 percent women, the mean for the gender variable should be 1.60. Similarly, if 85 percent of the cohort are white, the mean for Race: white should be 1.85. It should also be emphasized that the high school grade averages must be converted to the eight-point scale (see footnote "b" in the Tables) <u>before</u> the mean is calculated.

Evaluating Expected and Actual Retention Rates

Institutions that are highly successful at retaining their students should have actual retention rates that exceed their expected rates, whereas those institutions that have ineffective retention programs would be expected to have actual retention rates that fall substantially below their expected rates. Institutions with average retention capacity should have expected and actual rates that are very similar. While there are no hard-and-fast rules for deciding if expected and actual retention rates are essentially "the same," when the difference between these rates exceeds $\pm .10$, we are approaching a discrepancy which could be viewed as significant from both a practical as well as a statistical perspective (whether such a difference is indeed statistically significant would depend upon the size of the cohort being studied and the "p" value–.05, .01, etc.–that is, the amount of risk that the investigator is willing to take in inferring that the expected and actual rates are indeed different).

Recent research on retention suggests that there are a number of "environmental" factors that are known to influence an institution's actual retention rate, over and above the influence of student input characteristics (Astin, 1993b). One such factor is the students major field. Institutions enrolling many students in fields like business, psychology, or other social sciences would be expected to have higher-than-expected retention rates, or as those enrolling large numbers of students majoring in engineering would be expected to have lower-than-expected rates. The negative effect of engineering majors may well be an artifact, given that many engineering majors take longer than four years to complete a bachelor's degree.

Another major factor that increases these students' retention chances is living in a campus residence hall during the freshman year. Thus, institutions with required freshman residency or that house a large percentage of new students in campus residence halls would be expected to have higher-than-expected retention rates, whereas purely commuter institutions would be expected to have somewhat lower-than-expected rates. Institutional size, on the other hand, tends to have a negative effect on retention.

In short, institutions that are attempting to understand why their actual and expected retention rates may differ should keep these factors in mind. It is also important to realize that small size and residential facilities do not necessarily create actual retention rates that are higher-than-expected, nor does large size and a lack of residential facilities necessarily cause the institution's actual rate to be lower-than-expected. Rather, there are tendencies for size and residence to affect retention in the manner just described (Astin, 1993b).

References

Astin, A.W. "Assessment and Educational Reform: A Question of Values." Keynote Address presented at the 1993 Annual Meeting of the Association for Institutional Research, Chicago, May 17, 1993a.

______. What Matters in College: Four Critical Years Revisited. San Francisco: Jossey-Bass, 1993b.

______. Preventing Students from Dropping Out. San Francisco: Jossey-Bass, 1975.

Dey, E.L. and Astin, A.W. Predicting College Student Retention: Comparative Data from the 1982 Freshman Class. Los Angeles: UCLA Higher Education Institute, 1989.

Table 1

Predicting Bachelor's Degree Completion [a]
Using Different Combinations of Input Variables
(N=39,243)

Input Variable	b coefficient using formula 1	2	3	4
Average High School Grades [b]	.0915	.0622	.0562	.0563
SAT Math [c]		.000454	.000584	.000552
SAT Verbal [c]		.000433	.000404	.000376
Sex: Female [d]			.0741	.0736
Race: White [e]				.0250
Race: Native American [e]				-.1214
Race: African-American [e]				-.0560
Race: Chicano [e]				-.0709
Constant (a)	.0069	-.2729	-.4055	-.1671
Multiple R	.288	.327	.334	.338

[a] Within four years after entry as a full-time freshman (retained=1, not retained=2)
[b] A or A+=8, A-=7, B+=6, B=5, B-=4, C+=3, C or C-=2, D or less=1
[c] Includes ACT converted to SAT (see Appendix)
[d] Female=2, male=1
[e] Racial category=2, other=1

Table 2

Predicting Bachelor's Degree Completion or Four Years of Enrollment [a]
Using Different Combinations of Input Variables
(N=39,243)

Input Variable	b coefficient using formula 1	2	3	4
Average High School Grades [b]	.0810	.0551	.0517	.0515
SAT Math [c]		.000469	.000543	.000529
SAT Verbal [c]		.000302	.000285	.000278
Sex: Female [d]			.0421	.0420
Race: Native American [e]				-.1297
Race: African-American [e]				-.0299
Constant (a)	.1595	-.0846	-.1600	.0145
Multiple R	.264	.299	.302	.303

[a] Degree completion within four years after entry as a full-time freshman or four years of enrollment (retained=1, not retained=2)
[b] A or A+=8, A-=7, B+=6, B=5, B-=4, C+=3, C or C-=2, D or less=1
[c] Includes ACT converted to SAT (see Appendix)
[d] Female=2, male=1
[e] Racial category=2, other=1

Table 3

Predicting Bachelor's Degree Completion, Four Years of Enrollment, or Being Currently Enrolled [a] Using Different Combinations of Input Variables (N=39,243)

	b coefficient using formula			
Input Variable	1	2	3	4
Average High School Grades [b]	.0723	.0467	.0458	.0457
SAT Math [c]		.000492	.000512	.000502
SAT Verbal [c]		.000267	.000263	.000258
Sex: Female [d]			.0111	.0110
Race: Native American [e]				-.1070
Race: African-American [e]				-.0203
Constant (a)	.2685	.0269	.0070	.1445
Multiple R	.246	.285	.286	.287

[a] Degree completion within four years after entry as a full-time freshman, four years of enrollment, or being currently enrolled (retained=1, not retained=2)

[b] A or A+=8, A-=7, B+=6, B=5, B-=4, C+=3, C or C-=2, D or less=1

[c] Includes ACT converted to SAT (see Appendix)

[d] Female=2, male=1

[e] Racial category=2, other=1

Appendix
Converting ACT Scores to SAT Equivalents

The ACT equivalent was obtained by summing three ACT subtests (English, Sciences, Social Sciences) and converting to SAT equivalent by the equipercentile (N=14,865). The sum of the three (range 3–108) ACT subtests was used (rather than sir ACT English subtest) because it resulted in a better correlation with the SAT Verbal score (r r=.69). If a record had one or more of the ACT subtests missing, the entire record was from the file. The resulting conversion table is shown below.

ACT Sum	SAT Verbal	ACT Sum	SAT Verbal	ACT Sum	SAT V
108	800	72	500	36	30
107	800	71	490	35	30
106	800	70	480	34	29
105	800	69	480	33	28
104	800	68	470	32	28
103	800	67	460	31	27
102	800	66	460	30	26
101	790	65	450	29	26
100	770	64	440	28	25
99	760	63	440	27	25
98	750	62	440	26	240
97	740	61	430	25	230
96	730	60	430	24	230
95	720	59	420	23	220
94	710	58	420	22	220
93	700	57	410	21	210
92	690	56	410	20	210
91	680	55	400	19	210
90	670	54	400	18	210
89	660	53	390	17 or below	200
88	640	52	390		
87	630	51	380		
86	620	50	380		
85	610	49	370		
84	600	48	370		
83	590	47	360		
82	580	46	360		
81	570	45	350		
80	560	44	350		
79	550	43	340		
78	540	42	340		
77	540	41	330		
76	530	40	320		
75	520	39	310		
74	510	38	310		
73	510	37	310		

ACT equivalent is obtained by an equipercentile conversion of the ACT Mathematical subtest score (range 1–36) to SAT[1]. The correlation between SAT–M and converted ACT–M is .85 (N=14,000).

ACT Math	SAT Math
36	780
35	750
34	730
33	710
32	700
31	680
30	660
29	640
28	610
27	590
26	560
25	530
24	510
23	500
22	480
21	470
20	460
19	450
18	440
17	430
16	410
15	390
14	380
13	370
12	360
11	350
10	340
9	330
8	330
7	320
6	300
5	290
4	280
3	270
2	260
1	240

1 Adapted from Dey and Astin (1989).

analysis of a corollary of the Deviation Hypothesis. *Journal of Counseling Psychology*, 1967, 14, 462–472.
Guttman, L. To what extent do communalities reduce rank? *Psychometrika*, 1958, 23, 297–308.
Horst, P. *Factor analysis of data matrices*. New York: Holt, Rinehart, & Winston, 1965.
Horst, P. *Psychological measurement and prediction*. Belmont, Calif.: Wadsworth Publishing Co., 1966.
Kendall, M. G. and Stuart, A. *The advanced theory of statistics*. Vol. 1. London: Charles Griffin & Co., Ltd., 1969.
Loevinger, J. Objective tests as instruments of psychological theory. *Psychological Reports*, 1957, 3, 635–694.
Jackson, D. N. *Personality Research Form Manual*. Goshen, N. Y.: Research Psychologists Press, 1967.
Jackson, D. N. A sequential strategy for personality scale construction. In C. D. Spielberger (Ed.), *Current topics in clinical and community psychology*. Vol. 2. New York: Academic Press, 1970, pp. 61–96.
Jackson, D. N. The dynamics of structured personality tests: 1971. *Psychological Review*, 1971, 78, 229–248.
Neill, J. A. and Jackson, D. N. An evaluation of item selection strategies in personality scale construction. EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT, 1970, 30, 647–661.
Stricker, L. J. "Test wiseness" on personality scales. *Journal of Applied Psychology*, 1969, 53, (3, Pt. 2).

Fat #1 GPA #1=1

EDUCATIONAL AND PSYCHOLOGICAL MEASUREMENT
1976, 36, 135-140.

THE USE OF MULTIPLE REGRESSION ANALYSIS TO GENERATE CONDITIONAL PROBABILITIES ABOUT THE OCCURRENCE OF EVENTS

RICHARD F. HAASE
University of Massachusetts, Amherst

The use of Multiple Regression Analysis to compute a wide variety of statistics (ANOVA, Covariance, X^2, etc.) is becoming increasingly popular. An *MR* procedure for computing conditional probabilities of the type, $P(A/B) = P(B/A)P(A)/P(B)$, is reviewed. While it has not received wide attention in the psychological and educational literature, it has seen considerable application in the economic literature. The present paper emphasizes its potential usefulness in other areas of the behavioral sciences.

DURING the past five years educational and psychological research has seen a renewed interest in the statistical technique of Multiple Regression analysis (MR) for handling a vast array of data analytic problems (Darlington, 1968; Cohen, 1968; Hurst, 1970). Historically, there appeared to develop two camps regarding the appropriate data analytic technique for various problems. This seems to have resulted in a dichotomy roughly analogous to "experimental" versus "ex post facto" research. Due to what appears to this author as a rather unfortunate historical accident, analysis of variance techniques (ANOVA) and correlational techniques became associated respectively with "experimental" and "ex post facto" research paradigms.

I would submit that this state of affairs resulted largely from the principals involved in the development of each technique, that is, Fisher's analysis of variance being closely associated with his precise experimental work in agriculture, and correlational techniques becoming primarily associated with the ex post facto work in genetics of Galton and Pearson. As camps tend to do, they attract followers and

187952

self-perpetuate. Evidence that both camps have survived and are well can be found by examining a number of psychological and educational statistics texts which have been widely used throughout the past decade (Guilford, 1965; Hays, 1963; McNemar, 1962; Wert, Neidt, and Ahmann, 1954). It it noteworthy that these sources all treat MR and ANOVA extensively, but as separate topics. Little mention is made of their fundamental, underlying similarities.

What now appears to be re-appreciated is the fact that both ANOVA techniques and MR are part and parcel of the same domain, they are indeed fundamentally based in the general linear model,

$$Y = a + bX + e, \tag{1}$$

which submits that any response variate can be understood as a linear, additive combination of a constant (a), representing the value of y when $X = O$, one or more independent variables (X), appropriately weighted by some coefficient (b) derived from the data, and a source of random error (e). A more complete description of the general linear model can be found in Mendenhall (1968), Winer (1972) and Edwards (1968).

Perhaps more important to the behavioral scientist is the appearance of several references which have lucidly explicated the underlying similarities between *MR* and ANOVA and which have advocated *MR* techniques as appropriate for a wide range of data analytic problems (Cohen, 1968; Kelly, Beggs and McNeil, 1969; Kerlinger and Pedhazur, 1973). The purpose of this paper is to illustrate the use of *MR* for computing conditional probabilities of occurrence of events based on the functional relationship between a dependent response variate and one or more independent predictors. The conditional probabilities computed using an *MR* routine are patterned after those computed following Bayes' formula:

$$P(A; B) = \frac{P(B/A)P(A)}{P(B)} \tag{2}$$

The advantages in conceptualizing a functional relationship through *MR* as a series of conditional probability statements seem to be two: (1) it allows the researcher to make explicit probability statements about the occurrence of some dependent event as a function of one or more independent conditions. Such information is not available from an ordinary *MR* routine or from an ANOVA conceptualization of the problem; and (2) communication of results to the reader unfamiliar with the components of regression analysis is facilitated. While the consumer of educational and psychological research may not be familiar with regression analysis, (or statistical analysis in general), the

majority of that consumer group will have at least an intuitive grasp of the meaning of the probability of the occurrence of an event. The remainder of this paper is devoted to an explication of the set up of such an *MR* problem, and to the presentation of an example of its use.

Coding the Variables

The coding of the data matrices for the set up of the problem under discussion here is a straightforward coding of both the dependent criterion and the independent predictors as binary dummy variables. The particulars of dummy coding will not be reiterated since excellent discussions can be found in Cohen (1968) and Kerlinger and Pedhazur (1973).

If one's variables exist as continuously distributed data then one might recast the data as a dichotomy by splitting the distribution at the median and coding high and low values. Although treating continuous data in this fashion renders it slightly less sensitive it is necessary to do so to provide values of Y which are within the limits $(0 - 1)$ of probability statements. A discussion of some of the difficulties involved in categorizing continuous data can be found in Kerlinger and Pedhazur (1973).

Similarly for the predictor variables, they are cast in the form of dichotomous dummy variables. Once this has been accomplished the problem is solved in the usual *MR* fashion by solving for the appropriate unknowns in the model,

$$Y = a + b_1X_1 + \ldots + b_nX_n + e \tag{3}$$

where,

Y = a predicted value for the response variate,
a = a constant, the value of y when $X_1 \ldots X_n = 0$,
$b_1 \ldots b_n$ = the least squares regression coefficients,
$X_1 \ldots X_n$ = the n independent predictors,
e = errors of measurement.

An example of the use of *MR* in computing conditional probabilities is drawn from a study by Haase, Story, Bluestein, Slovin, Wolff and McEleney (1973). This project was concerned with assessing the relationships between perceptions of university residence halls as a function of sex and coeducational living. The example is drawn from that part of the study which assessed students' perception of the "ideal" residential environment. The instrument employed was the University Residential Environment Scale (URES) developed by Gerst and Moos (1972). The URES is composed of 10 subscales measuring a variety of aspects of the perceived environment. The scale reflecting orientation toward academic achievement is presented for illustration here.

The results of the regression analysis of perceived academic achievement regressed on the predictors of sex of subject (coded 1 = male, 0 = female) and types of house (coed = 1, single sex = 0) are presented in Table 1. The results are based on a solicited, nonrandom sample of 362 subjects.

The regression equation for assessing the conditional probability of perception of high levels of academic achievement given knowledge of sex of respondent and type of house then is,

$$Y = .5285 - .1946\,(\text{Sex}) + .0943\,(\text{House Type}). \tag{4}$$

In our example, academic achievement was coded 1 = high (above the median of the distribution) and 0 = low. Similarly the predictor variables can take on values of 1 or 0. Thus for sex, 1 = male, 0 = female; and for house type 1 = coed, 0 = single sex. Once the least squares regression coefficients are solved one can substitute appropriate values of the predictors into the equation (presence or absence) and solve the equation to yield the conditional probability of the perception of academic achievement given the four possible combinations of sex and house type. For example the probability that high levels of academic achievement will be perceived in a house given that the respondent is male and living in a single sex house is given by,

$$P\,(\text{acad. achiev./male, single sex}) = .5285 - .1946(1) + .0943(0). \tag{5}$$

$$\text{Probability} = .33. \tag{6}$$

In like fashion,

$$P\,(\text{acad. achiev./female, single sex}) = .53. \tag{7}$$

$$P\,(\text{acad. achiev./male, coed}) = .43. \tag{8}$$

$$P\,(\text{acad. achiev./female, coed}) = .62. \tag{9}$$

TABLE 1
Regression Analysis of Academic Achievement (URES) Predicted by Sex of Subject and Type of House

Dependent Variable	R	R²	F	P
Academic Achievement	.21	.04	8.12	<.01

Predictor Variables	Regression Coefficients	F	P
Sex	.1946	13.93	<.01
House type	.0943	3.01	<.10
Constant	.5285		

The four possible conditional probabilities computed from the regression equation (3) reveal an interesting pattern of predictions about the perception of academic achievement orientation in certain types of university residence halls. As is perhaps obvious male single-sex residences do not reflect especially good odds of demonstrating high levels of academic achievement. Conversely female, coed houses show a pattern of stronger proclivities toward academic pursuits.

The use of an *MR* model to achieve the results outlined above proved useful in the overall project from a variety of vantage points. It made quite explicit the prediction of the likelihood of satisfaction with residence halls based on the predictors employed. While we could have achieved a similar understanding of the interrelationships between the variables based only on the traditional *MR* results or an ANOVA conceptualization of the problem, the technique's greatest pragmatic usefulness came from the ease of communication of results to our consumers who have minimal understanding of the bases of statistical analysis.

While this technique has not appeared with any degree of frequency in the educational and psychological literature, it does seem to have a wider application in economics (see for example, Johnston, 1972; Orcutt, Greenberger, Korbel, and Rivlin, 1961). A particularly interesting example of its application to the study of litter control can be found in Finnie (1973).

This paper has attempted to illustrate a little used, but potentially useful, application of Multiple Regression analysis in behavioral research. In addition to its specific advantages it is hoped that this paper has also served to reinforce the notion that Multiple Regression analysis is a highly flexible data analytic strategy.

REFERENCES

Cohen, J. Multiple regression as a general data-analytic system. *Psychological Bulletin*, 1968, 70, 426–443.

Darlington, R. B. Multiple regression in psychological research and practice. *Psychological Bulletin*, 1968, 69, 161–182.

Edwards, A. L. *Experimental design in psychological research*. New York: Holt, Rinehart and Winston, (2nd ed.), 1959; (3rd ed.), 1968.

Finnie, W. C. Field experiments in litter control. *Environment and Behavior*, 1973, 5, 123–144.

Gerst, M. and Moos, R. The social ecology of university student residences. *Journal of Educational Psychology*, 1972, 63, 513–522.

Guilford, J. P. *Fundamentals of statistics in psychology and education*. New York: McGraw-Hill, 1965.

Haase, R. F., Story, R. W., Bluestein, D., Slovin, T., Wolff, T., and