An Analytic Model to Assist Academic Advisors

Stephen L. DesJardins, University of Michigan
Wang Jie, The University of Iowa

We detail how academic advisors at two land grand universities benefit from the identification of factors related to poor academic performance of first-year students. We used a multivariate statistical model and data from one institution to identify characteristics of students at-risk of earning low grade point averages. We showed through a second application of the statistical model that first-year dropout was directly related to grade point average.

KEY WORDS: at-risk, attrition, grade-point average, retention, student characteristics, tools for advising

Introduction

Effective academic advising is especially critical for first-year students, who are at the greatest risk of academic difficulty (Wolf & Johnson, 1995) and dropout (Pascarella, 1986; Pascarella & Terenzini, 1980). For instance, almost one half of the students who entered higher education left before completing either an associate or baccalaureate degree (Adelman, 1999; Gerald, 1992), and approximately 75% of nongraduates dropped out in their first year (Tinto, 1987). According to an ACT (2000) report, approximately one third of all students who matriculated to college in 1999 did not reenroll in the fall of 2000. This percentage is the lowest recorded since the inception of this ACT study in 1983.

Advising efforts have become increasingly important because a more diverse student body is enrolling in institutions of higher education (IHEs) (Zusman, 1994). Postsecondary students are likely to be first-generation college attendees, women, from minority groups, and older than students who have traditionally enrolled in college. These students may have special problems in adapting to college life, and academic advising may be a way to assist them in making the transition to college. Also, because of social and financial pressures, some institutions are admitting more students who may be academically ill prepared and therefore at risk of poor academic performance (Thombs, 1995). Thus, the identification of at-risk matriculating students is important to IHEs. Eno, McLaughlin, Sheldon, and Brozovsky (1999) demonstrated that statistical models can be used to predict entering students' freshman grades and that this information can be used to effectively advise students.

We discuss how personnel at two large IHEs identified the factors related to poor academic performance and used this information to change policies and processes related to academic advising. Although policy makers at these institutions had different reasons for addressing academic performance, they employed similar analytic strategies. Because these institutions are located in separate states, have unique problems, and serve students with different characteristics, administrators at each school employed separate statistical models. Nonetheless, administrators at each institution wanted to generate a list of student risk factors so that advisors could focus more of their limited time on advisees who were predicted to be at-risk.

To assist academic advisors, we developed a multivariate regression model to show that noncognitive factors can be used effectively to predict grade-point average (GPA) in first-year students at the University of Minnesota and the University of Iowa. Then, we estimated a second model that demonstrated that first-year GPA and other factors could be used to accurately predict year-one attrition. These two statistical models were estimated using only one half of the original sample. We applied the regression results to the other half of the sample and produced predicted GPA and retention values for each student in this "validation" sample. This approach allowed us to test the predictive accuracy of our statistical models by comparing predicted GPAs and retention rates to actual GPAs and retention rates available in the validation sample. The two statistical models accurately predicted GPA and dropout at the University of Minnesota and the University Iowa.

Background

In late 1998, the director of the University of Minnesota's College of Liberal Arts Undergraduate Advising Office shared that freshman enrollments at the university were increasing but the number of academic advisors available in the advising office had not increased at a concomitant rate. He stated that shortened advising sessions for entering freshmen was one solution under consideration. After a short discussion about how the advising process worked, we suggested building a statistical model that would identify the factors related to poor first-year academic performance and providing this information to advisors so that they could more efficiently target their limited advising time. We argued that with appropriate knowledge about at-risk factors, advisors may be able to make their efforts more productive. thereby allowing them to serve the additional students expected to matriculate the subsequent fall.

In the late 1990s, the University of Iowa retention rates of first-year students were stuck at approximately 83%. During this time, improvement of undergraduate retention and graduation rates was the top priority of the Iowa administration. Some decision makers within the institution felt that retention rates would improve if retention efforts were focused on freshman academic performance. They were considering an academic intervention that targeted at-risk and underperforming freshmen and looked to the Undergraduate Academic Advising Center to take the lead on this project. The Advising Center was constrained by a staff shortage, and it did not have (although administrators were considering) an appropriate intervention (e.g., a first-year seminar course for probationary students) for at-risk students. We suggested that the Iowa advising staff implement a statistical model similar to the one used at the University of Minnesota to help identify the factors related to academic problems in the first year. We pointed out that the analysis would not only provide institutional decision makers and academic advisors with statistical information that was previously unavailable, but the results of the analysis might be used to justify the expenditures needed to serve freshmen at risk for attrition.

The Literature

Many researchers have found that precollege performance and cognitive factors, including high school achievement and college entrance-exam scores, are correlated with poor academic performance in college and are also related to subsequent dropout (Pasearella & Terenzini, 1991). Ting (1997) noted that high school rank and ACT scores were effective in predicting first-semester GPA. Nisbet, Ruble, and Schurr (1982) found that high school rank percentile and SAT scores could be used to help predict student retention. Thombs (1995) found that statistical models could predict whether or not a student was likely to be put on probation. Thomhs (1995, p. 285) also found that a model including cognitive factors and variables related to problem behaviors "allowed for correct classification of 63.14% of the freshmen" from a public college in New York State.

Although researchers have often used cognitive factors to help explain student academic performance and dropout behavior, lately more research has been conducted on how noncognitive factors are related to student academic outcomes. For instance, Ting (1997) found that besides high school class rank and ACT exam scores. successful leadership experiences and demonstrated community service were effective predictors of academic success. Ting also found that demonstrated community service and preference for long-range goals were significant predictors of student retention. Other studies have also revealed that psychosocial predictors are useful in predicting college performance and retention (Fuertes. Sedlacek. & Liu, 1994: Tracey & Sedlacek, 1984. 1989).

Thombs (1995) also used noncognitive factors to distinguish between students on probation during the freshman year and students who were not on probation. Thombs found that in addition to cognitive factors, measures of study-habit problems, career-plan certainty, and time-management issues were useful in discriminating between freshmen on academic probation and those not on probation. Taking a page from Thombs's research, we included noncognitive measures in a model to be used to predict academic difficulty in college freshmen. Specifically, we used information regarding the level of student self-reported needs for assistance in a number of areas: study habits, reading, mathematics. writing, personal issues, and educationalioccupational plans. These factors may be associated with poor academic performance and might be used to improve statistical models designed to explain and predict freshman-year academic difficulty.

The lack of noncognitive information about students often generates a problem in statistical model development. However, a potential source of valuable attitudinal and other noncognitive information is often overlooked. When students take the ACT Assessment (or the SAT college-entrance exam), they fill out a questionnaire known as the ACT Student Profile Questionnaire (SPQ). This survey contains a plethora of cognitive and noncognitive information about test takers, such as the student's educational plans, interests, and needs; information about the student's demographic characteristics, background, and high school experiences; and data regarding the student's abilities, preferences, and needs. The responses from these surveys are routinely available to institutions and contain a very valuable but underutilized source of information. We demonstrate how coupling the SPQ information with institutional data can help IHE administrators develop models that will assist them in understanding better the factors that are related to poor academic performance during the freshman year.

First, we developed a multivariate regression model to show that noncognitive factors can be used effectively to predict GPA in first-year students at the University of Minnesota and the University of Iowa. Then we estimated a second model that demonstrated that GPA could be used as an accurate predictor of attrition. We built these models by using one half of the original sample (known as the "developmental" sample). Then, we used the statistical results to predict GPA and retention for the holdout sample. By comparing predicted GPAs and retention rates to actual GPAs and retention rates that were a part of the holdout sample, we were able to test the predictive accuracy of our statistical models. As we expected, the two statistical models accurately predicted GPA and dropout, respectively, at the University of Minnesota and the University Iowa.

Methodology

Participants

We used data from first-year students matriculating into the liberal arts colleges of two large land-grant universities: the University of Minnesota and the University of Iowa. We restricted the analysis to liberal arts students because they comprised the majority of entering freshmen on each campus and most undergraduate academic advising was focused on these students at both IHEs.

At Iowa, all freshmen entering the College of Liberal Arts in the fall semesters of 1997 and 1998 who filled out the SPQ were included in the analysis (N = 6,738 students). This group represented 93% of all freshmen entering the institution in these years. After deleting records with missing information, we had an effective sample size of 5,060 or 70% of the original sample. We used observable characteristics (e.g., racelethnicity, age, gender, and ACT test scores) to determine that the study sample was representative of freshmen entering the College of Liberal Arts at the University of Iowa. At Minnesota, we identified a sample of 4,252 students. As with the Iowa sample, we used observable factors to determine that the students sampled were similar to all students matriculating to the College of Liberal Arts in the fall semesters of 1995 and 1996.

The Predictors

We included three groups of explanatory variables in the statistical models (Table 1): a) demographic factors including gender and racelethnicity; b) precollege performance variables including ACT English and math scores, high school rank percentile, and whether the student met the institution's high school preparation requirements in English, math, social science, natural science, and foreign language; c) college educational needs including whether the student indicated needing help with writing, reading, mathematics, study skills, personal concerns, or with educational and occupational plans, the number of hours per week a student planned to work during the first year of college, and the position of the institution in the student's college-choice set. The background and cognitive variables were chosen based on previous research of at-risk students. We included the survey information related to student needs to test whether these student responses could be used to help predict poor academic performance during the first year of college. We included the choice variable as a proxy for institutional fit, which Tinto (1993) found to be an important component of student-environment fit.

The Criterion Variables

In this study, we examined two different academic outcomes of the first year: academic performance, as measured by first-semester/term GPA, and first-year dropout. At the time of the analysis, a University of Minnesota academic year consisted of three terms, so the GPA score for the Minnesota sample was the student's GPA in the first term of enrollment (Minnesota has since changed to a semester academic calendar). At the University of Iowa, an academic year is comprised of two semesters; thus the criterion used was a student's GPA during the first semester. A Minnesota or Iowa student is placed on academic probation when his or her first term/semester GPA falls (approximately) below 2.00. Because most other institutions hold a similar 2.00 threshold, we used a dichotomous dependent variable as an indicator of whether a student had an unsatisfactory first-term/semester GPA: Students with GPAs less than 2.00 were considered to have unsuccessful first terms/semesters.

As noted, many students who find themselves in academic difficulty during their freshman year fail to reenroll in their sophomore year. To test the impact of poor academic performance on reenrollment rates at the study institutions, we also created a dichotomous dependent variable where 1 indicated that the student did not reenroll in the fall of the sophomore year and 0 meant that the student enrolled. This indicator variable was used as the dependent variable in the dropout model.

The Statistical Models

As noted at the bottom of Table 2, the percent of students who had GPAs less than 2.00 and the percentage who dropped out in year one were all in the extreme lower end of the probability distribution. Therefore, logistic regression was preferable to linear (ordinary least squares) regression. We estimated two different logistic regressions. The first was defined as

$$log \frac{P_i}{1 - P_i} = \alpha + \beta_i X_i + \delta_i Y_i + \gamma_i Z_i + \varepsilon_i \quad \textbf{(1)}$$

where P_i is the probability that student i has a first-term/semester GPA less than 2.00; X, is a vector of personal and demographic characteristics; Y_i represents measures of precollege academic performance or achievement; Z_i represents student i's college educational needs; a, β_i, ti,, and γ_i are estimated coefficients; and ε_i represents a random error term that is logistically distributed (see Table 1 for the variables included as independent variables). The dependent variable is the logarithm of the odds that a particular student will have a GPA less than 2.00.

We estimated a second logistic regression

$$log \frac{P_i}{1 - P_,} - a + \beta_i X_i + \delta_i Y_i + \gamma_i Z_i + \&, \quad (2)$$

where P, indicated the probability that student i drops out within the first year; X_i is the same vector of personal and demographic characteristics noted in the GPA model; Y_i represents the same measures of precollege academic performance or achievement used in (1) but also includes the student's first semesterlterm GPA; Z_i represents the same college educational-needs variables used in (1); a, β_i, ti,, and γ_i are estimated coefficients; and ε_i represents a random error term that is logistically distributed. In (2) the dependent variable is the logarithm of the odds that a particular student will drop out within year one.

We estimated these two statistical models by using maximum-likelihood estimation.

Results

Descriptive Data

As described in Table 2, the raciallethnic distribution of students in the College of Liberal Arts at Iowa was 88.2% white, 2.4% Asian American, 2.0% African Americans, 2.1% Hispanic/Latino, and American Indians made up 0.5% of the sample. Male students comprised 40.3% of the incoming College of Liberal Arts freshmen; 13.4 (15.5)% of all matriculants had ACT English (math) scores one standard deviation or more below the mean; 30.0% of the incoming arts and science students indicated needing help with their study skills; 26.7 (21.6)% stated a need for help with mathematics (reading), and 14.2% indicated they would need assistance with writing. A total of 31.7% of the Iowa first-year students planned to work up to 20 hours per week, and 4.9% planned to work more than 20

Table 1 Definitions of the predictors

Predictors	Definition
Demographic Factors	
Gender	An indicator (dummy) variable equal to one if the student is male
Ethnicity	
White	The reference group, a dummy equal to one if the student is Caucasian
Native American	A dummy equal to one if the student is an American Indian
Asian American	A dummy equal to one if the student is an Asian American
African American	A dummy equal to one if the student is an African American
Hispanic	A dummy equal to one if the student is Hispanic
Other ethnicity	Did not respond to item/indicated a group not noted above
Enrollment 1997 (IA)	A dummy equal to one if the student matriculated to Iowa in 1997
Enrollment 1998 (IA)	A dummy equal to one if the student matriculated to Iowa in 1998
Enrollment 1994 (MN)	A dummy equal to one if the student matriculated to Minnesota in 1994
Enrollment 1995 (MN)	A dummy equal to one if the student matriculated to Minnesota in 1995
Speak English at home	A dummy equal to one if the student speaks English at home

Table 1 Definitions of the predictors (continued)

Predictors	Definition
Precollege Performance	
ACT English	A dummy equal to one if the student's ACT English score is more than one standard deviation below the average score of all matriculating students
ACT math	A dummy equal to one if the student's ACT math score is more than one standard deviation below the average score of all matriculating students
High school rank percentile below top quartile	Equal to one if the student's high school rank percentile is not in the top quartile
High school preparation requirements	
English preparation	A dummy equal to one if the student did *not* study English for 4 years or more
Math preparation	A dummy equal to one if the student did *not* study math for 3 years or more
Social science preparation	A dummy equal to one if the student did *not* study social science for 3 years or more
Nature science preparation	A dummy equal to one if the student did *not* study nature science for 3 years or more
Foreign language preparation	A dummy equal to one if the student did *not* study foreign language for 2 years or more
College Educational Needs	
Help with educational plans	A dummy equal to one if the student needs help with educational and occupational plans
Help with personal issues	A dummy equal to one if the student needs help with personal concerns
Help with math skills	A dummy equal to one if the student needs help in math
Help with reading skills	A dummy equal to one if the student needs help reading
Help writing skills	A dummy equal to one if the student needs help in writing
Help with study skills	A dummy equal to one if the student needs help in study skills
Work plans	
Not work/Work up to 10 hourstweek	A dummy equal to one if the student plans not to work or plans to work up to 10 hours per week
10–20 hours/week	A dummy equal to one if the student plans to work between 11 and 20 hours per week
21–30 hourstweek	A dummy equal to one if the student plans to work between 21 and 30 hours per week
More than 30 hourstweek	A dummy equal to one if the student plans to work more than 30 hours per week
First choice	A dummy equal to one if the university was the student's first choice
Not sure of major	A dummy equal to one if the student is not sure of his or her major in college
Dependent Variable	
GPA < 2.00	A dummy equal to one if the student's first year GPA is less than 2.00
Dropout	A dummy equal to one if the student did not enroll in the fall of the sophomore year

hours per week.

Of the Liberal Arts students at the University of Minnesota who were chosen for the sample, 84.5% were white; 8.0% were Asian American; 3.2% were African American; 2.1% were Hispanic/Latino; and American Indians made up 0.8% of the sample. Male students accounted for about 42.4% of the sampled matriculants; approximately 15% of the sampled students had ACT English and math scores one standard deviation or more below the mean; 29.2% of the sample stated a need for help with their study skills; 25.0 (21.1)% stated that they needed help with their mathematics (reading) skills; and 16.4% indicated that they would like to get help in improving their writing skills. Of the sample. 39.1% planned to work up to 20 hours per week, and 10.5% planned to work more than 20 hours per week.

Explaining First Term/Semester GPA

The results of the logistic regression analysis of students' first-semesteriterm GPA are presented in Table 3. We determined significance at $p < 0.01$ for all the items in the regression analysis. High school rank percentile, ACT English and math test scores, needing help with studying and writing in college, and planning to work up to 30 hours per week were statistically significant predictors of low GPA at both institutions. Regression techniques provide not only the direction of the relationship between the dependent and independent variables, but also an estimate of the magnitude of these relationships.

Status as an African American and needing help with educational and occupational plans were significant only in the Iowa sample (N = 5,060). An African American student was twice as likely as his or her white counterpart to have a GPA below 2.00. A student's inability to meet the high school natural science requirements was the only high-school requirement predictor variable significant in the Minnesota sample (N = 4,252). Students who did not meet the science requirement were 1.4 times more likely to earn a GPA less than 2.00 than were students who met the requirement.

Table 3 indicates that the odds of earning a first-term/semester GPA less than 2.00 varied by institution and for the explanatory variables included in the institutional models. For example, students with high school ranks below the 75th percentile were 2 times more likely (at Minnesota) and 4 times more likely (at Iowa) to be on probation than were students with high-school rank percentiles in the top quartile (the reference group). At Iowa and Minnesota, students who earned ACT English scores one standard deviation or more below the matriculating class average had 1.3 and 1.5 times greater odds of having GPAs less than 2.00 (respec-

Table 2 Descriptive statistics of participants

	University of Iowa		University of Minnesota	
Variable Category/Name	***n***	**%**	***n***	**%**
Demographic Factors				
White	5,943	88.2	3,594	84.5
American Indian	31	0.5	35	0.8
Hispanic	143	2.1	88	2.1
Asian American	164	2.4	341	8.0
African American	135	2.0	136	3.2
Other ethnicity	322	4.8	51	1.2
Male	2,718	40.3	1,804	42.4
Speak English at home	5,393	80.0	4,050	95.2
1998 Cohort	3,508	52.1	—	—
1995 Cohort	—	—	2,390	56.2
Precollege Performance				
ACT English > 1 ***SD*** below mean	900	13.4	657	15.5
ACT Math > 1 ***SD*** below mean	1,047	15.5	630	14.8
High school rank below top 25%	3,456	51.3	2,554	60.1
English preparation < 4 years	418	6.2	291	6.8
Math preparation < 3 years	133	2.0	180	4.2
Social science preparation < 2 years	465	6.9	136	3.2
Nature science preparation < 3 years	312	4.6	351	8.3
Foreign language preparation < 2 years	215	3.2	316	7.4

Table 2 Descriptive statistics of participants (continued)

	University of Iowa		University of Minnesota	
Variable Category/Name	***n***	**%**	***n***	**%**
College Educational Needs				
Educational Plans	2,951	43.8	2,046	48.1
Personal Plans	514	7.6	367	8.6
Mathematics	1,814	26.9	1,064	25.0
Reading	1,458	21.6	899	21.1
Writing	956	14.2	697	16.4
Studying	2,024	30.0	1,240	29.2
Work				
Not work or work up to 10 hours/week	3,245	48.2	1,213	28.5
10–20 hourslweek	2,137	31.7	1,664	39.1
20–30 hourslweek	331	4.9	447	10.5
More than 30 hourslweek	30	0.4	35	0.8
Not sure major	1,413	21.0	410	9.6
First choice	3,107	46.1	1,659	39.0
Dependent Variables				
GPA < 2.00	886	13.1	367	8.6
Dropout	1,120	16.6	752	17.7
Sample Size	6,738	—	4,252	—

Table 3 Multiplicative increase in the odds of first semester/term GPA < 2.00

	University of Iowa		University of Minnesota	
Variable Name (Reference Group)	**Odds Ratio**	**Coefficient Estimate**	**Odds Ratio**	**Coefficient Estimate**
High School Rank Percentile				
Below 75th percentile (top quartile)	4.0	1.386	2.0	0.693
ACT Math Score				
1 ***SD*** or more below class average (Not 1 ***SD*** below class average)	1.4	0.336	1.5	0.405
ACT English Score				
1 ***SD*** or more below class average (Not 1 ***SD*** below class average)	1.3	0.262	1.5	0.405
Indicated Needing Help With . . .				
Studying	1.7	0.531	1.7	0.531
Writing	1.4	0.336	1.3	0.262
Math	ns	—	ns	—
Reading	ns	—	ns	—
Educatiodoccupational plans (No indication of needing help)	0.79	–0.236	ns	—
Work				
10–20 hours/week	1.4	1.5	—	—
20–30 hours/week (Not work/work to 10 hourslweek)	1.8	1.7	—	—

Note. All results are significant at $p < 0.01$ level except where noted by ns.

tively) than their counterparts with higher ACT English test scores. Students who had ACT math scores one standard deviation or more below the matriculating class average ACT math score had 1.5 times (at Minnesota) and 1.4 times (at Iowa) greater odds of being on probation than were students with higher ACT math scores.

Students who planned to work were more likely to have GPAs less than 2.00. Students who worked 10 to 20 (20 to 30) hours per week were approximately 1.4 or 1.5 (1.8 or 1.7) times more likely to have GPAs less than 2.00 than were students in the reference group (those who did not work or who worked up to 10 hours per week) at Iowa and Minnesota. These magnitudes were invariant between study institutions.

A unique aspect of our study was the inclusion of factors related to the special educational needs of matriculating students. We found that at both institutions, students who expressed a need for help in developing their studying skills were about 1.7 times more likely to be put on probation than were students who did not express a desire for skill assistance. The importance of this variable in the statistical model is of particular interest: Needing help with studying is the second most powerful predictor (after high-school rank percentile) of being put on academic probation.

We also found that Minnesota and Iowa students who stated a need for help with writing were 30 and 40% (respectively) more likely to have GPAs less than 2.00 than were students who did not indicate a need for this type of assistance. Iowa students who expressed need for help with educational and occupational plans had lower odds of being on probation than their peers (79% that of students not indicating need for help in this area). The latter result is somewhat puzzling and deserves more investigation.

Because the results from these regression analyses suggest that noncognitive measures are predictive of poor academic performance, even when a host of demographic and cognitive measures are controlled, academic advisors have a new source of information to use when making assessments of students' likelihood of academic success. The consistency of the results between the two study institutions was unexpected and suggests that some common factors, at least at similar institutions, may help explain poor academic performance.

The Relationship Between First-Semester GPA and First-Year Dropout

We specifically focused on the relationship between a student's GPA in the first term/semester and his or her failure to reenroll in the fall of the sophomore year because advisors can use a history of poor performance early in a student's academic career as an indicator of future academic problems. However, even though we focused on the relationship between first semester/term GPA and dropout, the statistical model also controlled for the demographic, precollege performance, and college educational needs variables noted in Table 2. As noted in Table 4. our results indicate that one's first term/semester GPA is an effective and very powerful predictor of first-year dropout. As expected, the results indicate that the lower the GPA the higher the odds of dropout, and by using the regression model, we were able to put a magnitude on this relationship. Compared to students with first-term or semester GPAs greater than 3.00 (the reference group), students with GPAs between 2.00 and 3.00 had 1.5 and 1.2 times higher odds of dropping out of Minnesota and Iowa, respectively. For students with even lower GPAs, the odds of dropping out increase nonlinearly. Minnesota students with GPAs between 1.00 and 2.00 had odds of dropping out 5.8 times higher than their peers; those at Iowa were 12.0 times more likely to leave the institution than peers with higher grades. Students with GPAs less than 1.00 had high odds of dropping out in their first year. Minnesota (Iowa) students with grades in this range are 47.0 (687.0) times more likely than students with GPAs greater than 3.00 to quit school. The latter results suggest that students who have first-term/semester GPAs below 1.00 are almost certain not to be enrolled in the fall semester of their sophomore year.

Testing the Predictive Accuracy of the Models

Because the results of these models could be used to make decisions about academic advising, one needs to know the accuracy of the models for predicting student performance, as measured by GPA, or for predicting dropout before the sophomore year. To test the predictive efficacy of the models, one could categorize students at risk of probation or dropout, and then examine actual student performance in the future. However, to use this strategy the researcher must wait until the first semester is over to check the accuracy of the GPA model and must wait until fall of the sophomore year to assess the predictive accuracy of the dropout model.

Another strategy allowed us to immediately estimate the accuracy of the GPA and dropout models in predicting at risk behavior on subsequent enter-

Table 4 Multiplicative increase in the odds of first-year dropout

Variable Name (Reference Group)	University of Iowa Odds Ratio	University of Iowa Coefficient Estimate	University of Minnesota Odds Ratio	University of Minnesota Coefficient Estimate
GPA in First Semester/Term				
2.00 to 3.00	1.2	0.157	1.5	0.405
1.00 to 2.00	12.0	2.485	5.8	1.758
Less than 1.00	687.0	6.532	47.0	3.850
(3.00 or higher)				

Note. All results significant at $p < 0.01$ level.

ing cohorts. The original data on which the models were estimated are historical; therefore, we know the actual GPA and enrollment status of each individual in the data set. To test the efficacy of the GPA and dropout models, for each institution we randomly split the original data set into a developmental sample and a validation sample.' For each institution we ran two regressions on the developmental sample, one estimating first semesterlterm GPA, the other estimating the probability of first year dropout. Once we had the regression data, we applied the statistical regression formulas to the validation sample (by institution) to predict the first semesterlterm GPAs and dropout rates of the students in the validation group. Because the validation sample was constructed from historical data, we know each student's first semesterlterm GPA and whether or not they dropped out within the first year. Our application of the regression formula to the validation sample produced a prediction of first semesterlterm GPA and dropout for each student in the validation sample. The validation sample contained the actual and predicted values for both criterion variables (GPA and dropout), thereby allowing us to assess the predictive accuracy of our two statistical models.

Although several alternatives to test the predictive accuracy of these models are available (DesJardins, 2002), we chose the method used by Hosmer and Lemeshow (1989) and Lemeshow and LeGall (1994); this method is commonly available in most statistical software packages. Once we entered the data into the computer, the statistical software procedure regroups the validation data into 10 nearly equal-size groups or deciles (see column 2, Table 5).[2] Then the Hosmer-Lemshow Goodness-of-Fit Test (1989, pp. 140-45) is used to statistically test the within-decile accuracy of the estimated model. A significant HL Goodness of Test result provides evidence that the model does not fit the data, but a model that closely fits the data will produce an insignificant HL test.

Table 5 provides a great deal of detail about the accuracy of the GPA- and dropout-prediction models (to conserve on space, only the Minnesota results are displayed). Column one contains the decile groupings, which we formed by ordering on the predicted probabilities that a student was (was not) in the less-than-2.00 GPA (greater-than- or equal-to-2.00 GPA) category and then dividing all the 2,126 students in the validation sample into 10 (roughly) equal-size groups (see column 2). Students with low probabilities of having first-term GPAs less than 2.00 comprise the low-numbered decile groups, and students with high predicted probabilities of being on probation are in the higher numbered deciles. One can see that the accuracy of the predictions vary across deciles. In group (or decile) 1, the actual and predicted results are nearly identical. Five (200) students had GPAs less than 2.00, and the model predicted that 5.2 (199.8) would have (would not have) low GPAs. These results suggest that the model can be used to accurately predict the outcomes for students who have low probabilities of poor academic performance. But the model is less accurate when predicting students in the higher decile groups. For instance, in Group 10, 45 (167) students had (did not have) GPAs less than 2.00, and the model predicted that 51.2 (160.8) would have (would not have) low GPAs. Although one can see that Table 5 provides some evidence that the model can be

[1] Researchers could also use one cohort of historical data as the developmental sample (e.g., 1995 Minnesota matriculants) and then validate the model on a more recent cohort (e.g., all 1996 Minnesota matriculants).

[2] These tables are easily produced when using the logistic regression procedures in many mainstream statistical packages such as SAS, SPSS, and STATA (see the documentation for the LACKFIT option in SAS, the Hosmer-Lemeshow goodness-of-fit option in SPSS, and the LFIT option in STATA).

Table 5 Testing the predictive accuracy of the Minnesota GPA model

		Grade Point Average			
		Less Than 2.00		Greater Than or Equal to 2.00	
Decile Group	Total Number of Students	Actual *n*	Predicted *n*	Actual *n*	Predicted *n*
1	205.0	5.0	5.2	200.0	199.8
2	213.0	5.0	7.0	208.0	206.1
3	213.0	11.0	8.5	202.0	204.5
4	214.0	15.0	10.5	199.0	203.5
5	214.0	18.0	12.8	196.0	201.2
6	213.0	15.0	15.4	198.0	197.6
7	214.0	16.0	18.4	198.0	195.6
8	214.0	26.0	23.1	188.0	190.9
9	214.0	26.0	31.1	188.0	182.9
10	212.0	45.0	51.2	167.0	160.8
Total	2,126.0	182.0	183.2	1,944.0	1,942.8

Hosmer & Lemeshow Goodness of Fit Test Statistics

Chi-square	*df*	*pr* > Chi-square
8.2946	8	0.4052

used with accurate results, the hypothesis-based statistical evidence suggests that the model fits the validation sample. The HL goodness-of-fit statistic, displayed at the bottom of Table 5, is not significant at any conventional levels ($p = 0.4052$); therefore, the model can be used to predict likelihood of low GPAs among incoming Liberal Arts students at the University of Minnesota.

Table 6 provides information about the predictive accuracy of the model with regard to dropout at the University of Minnesota. Through observations of the table and statistical evidence (HL statistic $p = 0.4642$), one can see that the dropout model can also be used to accurately predict the behavior of students at the University of Minnesota.

We conducted similar analyses to test the predictive efficacy of the GPA and dropout models at the University of Iowa. The results indicated that the Iowa models were also effective in predicting the GPA and dropout outcomes.

Discussion

Because first-year students with low GPAs have relatively high odds of dropping out of IHEs, academic advisors can benefit from information about students who are at risk of poor academic performance. For instance, at Minnesota, administrators used the results of the regression models to develop a list of risk factors associated with poor first-year academic performance. They gave academic advisors a one-page list of risk factors associated with GPAs less than 2.00 and offered instructions on interpreting the magnitude associated with each variable (the relevant odds ratios). They told advisors that the more risk factors a student presented the more likely the student would be a poor academic performer. Advisors were instructed to use this information, and their professional judgment, to target their time toward at-risk students. No specific plan was implemented on how advisors should allocate their scarce advising time; the individual advisor could use the data at his or her discretion.

To be effective, Minnesota advisors needed more information than a list of advisee risk factors. They also needed to know how each of their advisees responded to the SPQ questions that were used as independent variables in the statistical model. For example, each advisor needed to know which students reported needing help with study habits. To provide this information to academic advisors, some institutional reporting procedures at the University of Minnesota were changed.

Prior to the initial advising appointment, academic advisors received a form called the Student Advising Profile (SAP). This form contained student information that advisors could use during initial and subsequent advising sessions. Because some of the risk factors identified by the statistical model were not previously provided on the SAP (e.g., student indications that studying assistance was needed), this information had to be added to the forms.

Table 6 Testing the predictive accuracy of the Minnesota dropout model

		Dropout Within the First Year of Enrollment		Retained to the Second Fall Term	
Decile Group	**Total Number of Students**	**Actual n**	**Predicted n**	**Actual n**	**Predicted n**
1	201.0	11.0	11.6	190.0	189.4
2	201.0	18.0	16.6	183.0	184.4
3	202.0	20.0	20.0	182.0	182.0
4	200.0	16.0	22.8	184.0	177.2
5	201.0	27.0	26.0	174.0	175.0
6	200.0	32.0	29.1	168.0	170.9
7	200.0	35.0	33.3	165.0	166.7
8	200.0	32.0	39.7	168.0	160.3
9	200.0	49.0	55.0	151.0	145.0
10	193.0	93.0	102.8	100.0	90.2
Total	1,998.0	333.0	356.9	1,665.0	1,641.1

Hosmer & Lemeshow Goodness of Fit Test Statistics

Chi-square	*df*	*pr* > Chi-square
7.6908	8	0.4642

One year after receiving the list of risk factors and SPQ information, as included in the adjusted SAP forms, advisors were actively using the risk factor lists and SAP information to help them in their advising efforts. The director of the office reported that advisors seemed willing to use this statistical information because it was easy to understand and it confirmed the "mental models" thev had generated about the factors related to poor academic performance. Advisors indicated that by coupling the results of the model with their professional experience they were more able to quickly identify at-risk students and to carefully scrutinize these students' academic plans.

The experienced academic advisors felt that the statistical model helped new and temporary academic advisors become more efficient more quickly than they would have been without the statistical results. Instead of needing years of experience to confirm their own instincts, the new advisors could quickly corroborate the statistical evidence and their own mental models regarding factors associated with unsatisfactory academic performance.

At Iowa. the results of the statistical analvsis were presented to decision makers in the academic advising office. Although Iowa advisors also had mental models that were consistent with the study results, the administrators had not known the magnitude of the relationships between at-risk factors and poor academic performance. Thus, the results from our studies provided useful information on specific characteristics that lead to unsatisfactory student performance and attrition.

Even though advising professionals at Iowa found the results of the model interesting and useful, they were dismayed that few intervention options were presented to help academically at-risk students. In the fall of 2000, they continued discussions about the results of the statistical analysis and considered a new course for at-risk students. The College Success Seminar (CSS) was offered to help at-risk students develop the skills, strategies, and habits that are essential for success in college. This voluntary, one-semester-hourcourse was instituted in the spring of 2001. It was immediately popular and seven sections of the course were full of students (eight in fall of 2002). Evidence available as of this writing suggests that more probationary students returned for a second year than before the addition of the CSS.

By adding a first-year experience course, advisors at the University of Iowa also focused efforts on students' early social and academic issues. This course was pilot tested during the fall semester of 2001 and as of the fall 2002 semester, 24 sections were offered for two semester credit hours. Although no formal evaluation of the course has been completed to date, a review of the course evaluations suggests that students find the course helpful: Approximately 95% of students filling out course evaluations indicated they would suggest the course to their peers.

At Iowa, administrators discussed the statistical model and sought to centralize data, such as the

ACT SPQ, for the advising unit to use. This operationalization process is continuing and the SPQ information will likely be available as will information provided by a pilot study being conducted by ACT. Thus, the analytic process, initiated by results of the regression model, not only provided valuable statistical information to advisors, but reinforced ongoing discussions about serving at-risk students at Iowa.

Although the results of our study may be specific to these two large public institutions, the process by which these studies were undertaken should be valuable to administrators at other IHEs as they develop an approach to improve academic advising efforts with limited resources. The quality of academic advising is positively related to GPA (Braxton, Duster, & Pascarella, 1988; Metzner, 1989) and student satisfaction (Gardner & Jewler, 1992) and negatively related to intent to leave the institution (Metzner, 1989). Academic advisors who used the information indicated that this analytic approach had helped to make academic advising more effective, and they noted that inexperienced advisors were particular beneficiaries. Although difficult to document the actual effects, the regression-model information about students seems to have helped advisors use their limited advising time more efficiently. The discussions that ensued while undertaking these studies also caused administrators to reevaluate institutional data collection and dissemination procedures. The analytic process supported efforts to change how IHEs service freshmen, especially those who have academic difficulty. Thus, the statistical-model approach appears to have met its initial objective to provide better information to advisors, but the process of self-inquiry also produced substantial unintended benefits.

References

ACT, Inc. (2000). *National dropout rates: Freshman to sophomore year by type ofinstitution.* lowa City, IA: Author.

Adelman, C. (1999). *Answers in the tool box: Academic intensity, attendance patterns, and bachelor's degree attainment.* Jessup, MD: U.S. Department of Education.

Braxton, J. M., Duster, M., & Pascarella, E. T. (1988). Causal modeling and path analysis: An introduction and an illustration in student attrition research. *Journal of College Student Development, 29*(3), 263–72

DesJardins, S. L. (2002). An analytic strategy to assist institutional recruitment and marketing efforts. *Research in Higher Education, 43*(5), 531–53.

Eno, D., McLaughlin, G. W., Sheldon, P., & Brozovsky, P. (1999). Predicting freshman success based on high school record and other measures. *AIR Professional File, 72*(Summer), 1–12.

Fuertes, J. N., Sedlacek, W. E., & Liu, W. M. (1994). Using the SAT and noncognitive variables to predict the grades and retention of Asian American university students. *Measurement and Evaluation in Counseling and Development, 27*(2), 74–84.

Gardner, J. N., & Jewler, A. J. (1992). Major, career, and academic advisors: Getting the most from your academic advisor. In J. N. Gardner & A. J. Jewler (Eds.), *Your college experience: Strategies for success.* Belmont, CA: Wadsworth Publishing Co.

Gerald, D. E. (1992). *Projections of education statistics to* 2003. Washington, DC: National Center for Education Statistics.

Hood, D. W. (1992). Academic and noncognitive factors affecting the retention of black men at a predominantly white university. *Journal* of *Negro Education, 61*(1), 12–23.

Hosmer, D. W., & Lemeshow, S. (1989). *Applied regression analysis.* New York: John Wiley and Sons.

Houston, L. N. (1980). Predicting academic achievement among specially admitted black female college students. *Educational and Psychological Measurement, 40*(4), 1189–95.

Lemeshow, S., & LeGall, J. R. (1994). Modeling the severity of illness of ICU patients: A systems update. *Journal of' the American Medical Association. 272*(13), 1049–55.

Metzner, B. S. (1989). Perceived quality of academic advising: The effect on freshman attrition. *Americun Educational Research Journal, 26*(3), 422–42.

Nisbet, J., Ruble, V. E., & Schurr, K. T. (1982). Predictors of academic success with high-risk college students. *Journul of College Student Personnel, 23*(3), 227–35.

Pascarella, E. T. (1986). A program for research and policy development on student persistence at the institutional level. *Journal of College Student Personnel, 27*(2), 100–107.

Pascarella, E. T., & Terenzini, P. T. (1980). Predicting freshman persistence and voluntary dropout decisions from a theoretical model. *Journal of Higher Education. 51*(1), 60–75.

Pascarella, E. T., & Terenzini. P. T. (1991). *How college affects students: Findings* and *insights from* *twenty years of research.* San Francisco: Jossey-Bass.

Poindexter, J. P., & Robinson, T. L. (1997). Relationships among racial identity attitudes, womanist identity, and self-esteem among African American college women. *Journal of College Student Development, 38*(3), 288–96.

Thombs, D. L. (1995). Problem behavior and academic achievement among first-semester college freshman. *Journal of College Student Development, 36*(3), 280–88.

Ting, S. R. (1997). Estimating academic success in the 1st year of college for specially admitted white students: A model combining cognitive and psychosocial predictors. *Journal of College Student Development, 38*(4), 401–409.

Ting, S. R., & Robinson, T. L. (1998). First-year academic success: A prediction combining cognitive and psychosocial variables for Caucasian and African American students. *Journal of College Student Development, 39*(6), 599–610.

Tinto, V. (1987). *Leaving college: Rethinking the causes and cures of student attrition.* Chicago: University of Chicago Press.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition* (2nd ed.). Chicago: University of Chicago Press.

Tracey, T. J., & Sedlacek, W. E. (1984). Noncognitive variables in predicting academic success by race. *Measurement and Evaluation in Guidance, 16*(4), 171–78.

Tracey, T. J., & Sedlacek, W. E. (1989). Factor structure of the non-cognitive questionnaire—Revised across samples of black and white college students. *Educational and Psychological Measurement, 49*(3), 637–48.

White, T. J., & Sedlacek, W. E. (1986). Non-cognitive predictors: Grades and retention of specially admitted students. *Journal of College Admission, 3*(Spring), 20–23.

Wolf, R. N., & Johnson, S. D. (1995). Personality as a predictor of college performance. *Educational and Psychological Measurement, 55*(2), 177–85.

Zusman, A. (1994). Current and emerging issues facing higher education in the United States. In P. G. Altbach, R. O. Berdahl, & P.J. Gumport (Eds.), *Higher education in American society.* Amherst, NY: Prometheus.

Authors' Note

Stephen L. DesJardins is Associate Professor; Center for the Study of Higher and Postsecondary Education, College of Education, at the University of Michigan. His research interests include strategic enrollment management issues, the study of student departure from college, policy analysis, and the economics of higher education. His work in these areas has been published in the Journal of Human Resources, Journal of Higher Education, Economics of Education Review, Research in Higher Education, Journal of College Student Retention, Journal of Student Financial Aid, *and* Higher Education: Handbook of Theory and Research.

Wang Jie is a doctoral candidate in the Educational Policy and Leadership Studies program in the College of Education, at the University of Iowa. Her research interests include student departure, management information, and student development.