How to Make Better College Admission Decisions:

Considering High School Quality and Other Factors

by Mo-Yin S. Tam and Uday Sukhatme

Abstract

A student's standardized examination score and high school percentile rank are two quantitative factors admission often uses to form a selection index, which is one of the main criteria used to make college admission decisions, especially in public universities. While the standardized scores are based on national examinations, the high school percentile rank is dependent on the quality of the student's high school. We constructed a modified high school percentile rank variable, to take into account both high school percentile rank and the quality of the student's high school. Using student data from the University of Illinois at Chicago, we showed a student's modified high school percentile rank as a substantially better indicator for university success than other commonly used variables. Based on this result, we recommend the use of a different admission selection index; one based on a student's standardized examination score and modified high school percentile rank. This new admission index should be used, together with other considerations such as diversity, to make better college admission decisions.

Mo-Yin S. Tam has been an Associate Vice Chancellor for Academic Affairs at the University of Illinois at Chicago (UIC) since August 2000. She is also a professor of economics at UIC. Dr. Tam is the PI of a recent NSF grant. In addition to examining various factors influencing college student performance, her recent research interests include the study of gender and ethnic gaps in information technology as well as the impact of diversity on student performance.

Uday Sukhatme has been the Dean of the College of Arts and Sciences at the State University of New York at Buffalo, since August 2002. Prior, Dr. Sukhatme served at the University of Illinois at Chicago for 22 years, where he was Head of the Department of Physics for seven years and vice provost for academic affairs for two years. Dr. Sukhatme's research interests are theoretical high-energy physics and quantum mechanics. He has authored more than 170 pieces for physics journals. Dr. Sukhatme has initiated numerous educational outreach programs aimed at high school students. He strongly believes that a careful analysis of various available input ingredients, when hiring faculty and admitting students, plays an important role in determining subsequent faculty retention and student success.

Introduction
Student applications for college admission contain many qualitative and quantitative attributes (input variables), which characterize the students' backgrounds and academic careers. These attributes often include name, address, date and place of birth, sex, ethnicity, high school name, standardized test score (on the ACT or SAT examinations), high school transcript, high school percentile rank, student essay, letters of recommendation, and description of extra-curricular activities. Admission decisions at the freshman level at universities are based on an evaluation of these attributes. Admission officers must ascertain which attributes, singly and in combination, provide the best indicators of subsequent university success, in order to make optimal and fair admission decisions, particularly at urban, public universities. With the severe constraints on capacity prevalent at these institutions, admitting students who do not graduate wastes limited educational resources partially subsidized by tax revenue. Diminishing state funding and increasing demand for accountability create considerable pressure for higher education to find new ways of assessing high school students' fitness for college study (Stern and Briggs, 2001).

We analyzed data from the University of Illinois at Chicago (UIC), where admission decisions were primarily based on a selection index, created using a student's ACT score and high school percentile rank. While ACT scores are comparable since all students take the same national examination, high school percentile ranks are dependent on the corresponding high school quality. A student has more difficulty achieving a high percentile rank in a more competitive high school. High school quality therefore should be taken into consideration in the formulation of the selection index.

The quality of a high school can be measured by a number of variables, including the academic performance of its students, the percentage of high school students going to college, the qualification of its teachers, the availability of advanced placement courses and the variety of courses offered. Due to limitation of data, we used the average ACT score of all students from the high school as a measure of its quality, recognizing that it did not fully capture the high school's quality.

We studied the student pipeline at UIC over a period of six years and found that much better admission decisions would have resulted from using the quality of the high school, as measured by the average ACT score of that high school, as one of the attributes. Specifically, we defined a new quantity, called the "modified student high school percentile rank," which incorporated a measure of high school quality. We found that the modified student high school percentile rank led to a new admission selection index capable of yielding much better admission decisions. Although we drew this surprising conclusion based on UIC data, it should be applicable to other urban public universities.

The specific procedure we developed did not make a priori guesses regarding the relative weights placed on different variables in making admission decisions. Instead, the importance was determined from a statistical analysis of previous data on successful students.

Description of the Study Sample and Student Pipeline
The sample for our study was the cohort of 2,529 freshmen who entered UIC in Fall 1994—1,402 males and 1,127 females. They had an overall average ACT score of 20.8 and an overall high school percentile rank of 73.0. We tracked the progress of this cohort for six years. Our motivation was to better understand the attributes of both successful and unsuccessful students, where success was measured by graduation within six years of starting studies at UIC.

Table 1 shows as a function of time, the percentages of students who continued to remain enrolled at UIC, those who graduated, those who dropped out in poor academic standing,

Table 1

Status of the UIC Fall 1994 cohort of freshmen as a function of time. The numbers give percentages of students at the end of an academic year, with year 1 denoting AY 1994-95.

At the end of academic year	0	1	2	3	4	5	6
Enrolled students	100	72	58	51	39	17	5
Graduated students	0	0	0	0	8	26	36
Dropouts in poor standing	0	22	32	36	39	41	42
Dropouts in good standing	0	6	10	13	14	16	17
Total	100	100	100	100	100	100	100

and those who left in good academic standing. At UIC, grades were given on a five-point scale, and students who left in good standing had a cumulative GPA≥3 and students who left in poor standing had a cumulative GPA<3. The category of students who dropped out in poor standing included both students who dropped out while on probationary status and those who the university dropped for unsatisfactory academic performance. We examined the roster of student identifiers annually and compared them with the starting 1994 roster. Thus, we included, in the graduating students category, those who were continuously enrolled and those who took a few intermediate semesters off from studies.

For the cohort we studied, the six-year graduation rate is 36 percent. Most of the attrition occurred in the first two years, during which 76 percent of students in poor standing dropped out. This highlighted the importance of retention programs in the early years for students at risk.[1] The four-year graduation rate was eight percent, though most of the students who were still enrolled at the end of four years did eventually graduate. This pattern supported the use of a six-year graduation rate as a better outcome measure for success than the classical four-year graduation rate (as used in the *U.S. News and World Report*, 2001).

Correlating Graduation Rate with Input Student Attributes

To determine the six-year graduation rate related to various quantitative attributes that characterized the students when they started their undergraduate studies at UIC, we considered four different input variables (attributes), one at a time. The four input variables were high school percentile rank, ACT score, the average ACT score of high school, and modified high school percentile rank. The first two variables, as indicated above, were used by UIC in formulating its admission selection index. The average ACT score of a high school measured the high school's quality. The modified high school percentile rank variable considered both the student's high school rank and the high school quality.

Our general approach consisted of plotting the data of the six-year graduation rate G_6 against the above four input variables (attributes), one at a time. The first plot, for example, was relating the six-year graduation rate to the high school percentile rank. We also constructed a straight line best fit to the plot of the data.[2]

We then considered the slope of the fitted straight line. Large positive (statistically significant) slopes corresponded to important variables in admission decisions, whereas zero or small or (statistically insignificant) slopes were indicators of unimportant attributes. For example, a plot of the graduation rate G_6 versus the number of the student's month of birth [1 for January, 2 for February, etc.] gave a flat straight line of essentially zero slope, indicating that the month of birth does not have any effect in predicting academic success—a predictable outcome!

A plot of G_6 versus the first input variable HSPR (a student's high school percentile rank) showed a small (statistically significant) slope of 0.004. In this determination, we noted that the range of values of the high school percentile rank is from 0 to 100.

Similarly, a plot of the data for G_6 versus the student's ACT score showed a positive correlation and statistically significant slope. However, in order to make a proper slope comparison between HSPR and ACT, we needed to modify the variable ACT so that it had a range of value similar to that of the high school percentile rank (from 0 to 100). The modified variable is denoted as NACT.[3] The slope of the graduation rate G_6 versus the student's NACT score is 0.008. Since the slope for high school percentile rank is, as indicated above, 0.004. Our results indicated that the impact of ACT score on G_6 was twice as large as the impact of the high school percentile rank on the graduation rate.

We then plotted G_6 versus the third input variable: the average high school ACT score (HSACT). We normalized this variable in the similar process as described above, so that the values of the normalized average high school ACT score (NHSACT) ranged from 0 to 100. Again, this allowed slope value comparison. Here the slope was found to be 0.010. Comparing this value to the slope of HSPR (0.004) and the slope of ACT (0.008) indicated, perhaps surprisingly, that, of the variables considered (ACT, high school percentile rank and average high school ACT), the average high school ACT variable (NHSACT) was the most significant for determining student success. Our results are listed in Table 2 and shown in Figure 1.

However, this new variable was not student-specific, because all students from the same high school have the same average high school ACT score. We used this result to construct a student-specific input variable.

Table 2

Least Squares Estimates of the Slopes of Various Input Variables versus the Six-Year Graduation Rate G_6 using the UIC Fall 1994 cohort of freshmen data

HSPR: High School Percentile Rank	0.004
NACT: Normalized ACT	0.008
NHSACT: Mormalized High School ACT	0.010
HRPR*: Modified High School Percentile Rank	0.015
Normalized SI: Selection Index using a linear combination of HSPR and NACT	0.011
Normalized SI*: Selection Index using a linear combination of HSPR* and NACT	0.016

Input variables are normalized where necessary so that they all have values between 0 and 100 for slope comparison.

Improved Student Attribute

Using the above results, we constructed an improved student attribute, which was a better indicator of student success. The new variable was a modified high school percentile rank (HSPR* defined by HSPR*=HSPRxHSACT/36). The new variable HSPR* took into account both a student's rank in the high school (HSPR) and a quantitative measure of the school quality (HSACT), and was also student-specific.

We examined the impact of this new variable on the six-year graduation rate using the same process described above. Since the values of the new variable ranged from 0 to 100, we could compare the value of the slope for this variable with that of the previously considered variables.

The slope of the straight line correlation between the graduation rate G_6 and HSPR* exhibited a (statistically significant) slope of 0.015. It was almost three times larger than the slope of the standard HSPR (which, is indicated earlier, is 0.004). Hence, the modified high school percentile rank was a much more improved variable than the standard high school percentile rank in determining eventual graduation.

Improved Selection Index

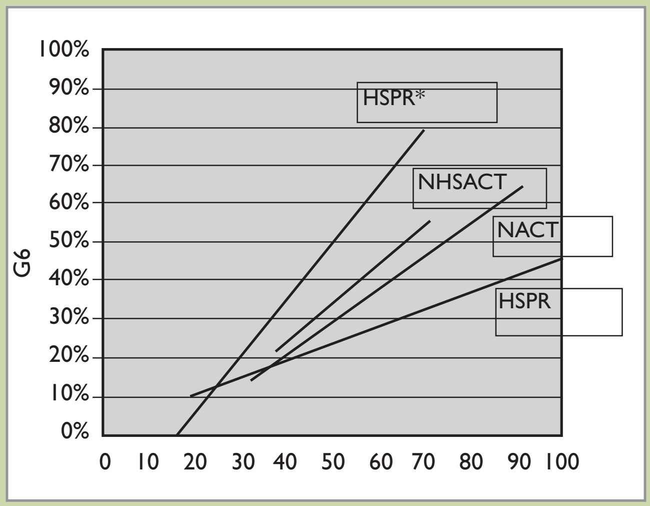
For admission purposes, many universities use a selection index, which is a linear combination of several variables with varying weights. The current UIC selection index is a linear combination of the student's HSPR and ACT, the weights being chosen through a process similar to the one described above.

In the previous discussion, we fit a straight line relating the six-year graduation rate to input variables (attributes), one at a time. Since the selection index is a combination of two variables (HSPR and ACT), we fit a straight line to the plot of six-year graduation rate to both of the variables.[4]

For slope comparison, we again used a normalized selection index variable so its values were between 0 and 100.[5] Using the student data of UIC, we found that the value for the slope of normalized selection index SI is 0.011, greater than either the slope of the high school percentile rank (0.004) and that of the ACT (0.008), but not as good as our newly defined variable HSPR*, the modified high school percentile rank (0.015).

Figure 1

The six-year graduation rate as a function of four normalized input variables. The normalized slopes plotted in the figure are: 0.004 (HSPR), 0.008 (NACT), 0.010 (NHSACT), 0.015 (HSPR*).



Based on our analysis, we recommended considering a better selection index, SI*, which was a linear combination of ACT and HSPR* with weights in the ratio again chosen so as to maximize the slope.[6] Indeed, the slope of SI* is 0.016, making it the best indicator of eventual graduation at UIC, compared to all the variables discussed in this study. Based on this result, we recommended consideration of the use of SI*, together with other considerations, such as diversity, in college admission decisions.

In addition to the overall results presented above, we have also examined how student success depends on gender [male, female] and race/ethnicity [Latino, Black, Asian, White]. We refer the interested reader to Tam and Sukhatme, 2003.

Discussion

Although the above UIC example makes use of just two variables (ACT and HSPR*), the discussion can be easily extended to more variables if desired. This may be necessary at other universities, depending on their character and location. The specific procedure we suggest for taking N available quantitative student-specific input variables into account and making better admission decisions is as follows:

1. Take each variable and normalize it if necessary, so its value is in the range 0 to 100 for slope comparisons.
2. Plot the data for the six-year graduation rate G_6 versus each of the single variables in turn, and make a straight line fit of the data points. Measure the slopes of each of these single variable correlation lines—they determine the relative significance of the impacts of various variables on the six-year graduation rate.
3. Compose a selection index (SI) as a linear combination of any number of these variables, choosing the weights of these variables to maximize the value of the slope of the corresponding SI. The construction procedure guarantees that SI also lies in the range 0 to 100, as desired. The weights of various variables are proportional to the slopes. Use the SI for making admission decisions.

While graduation is the ultimate test of success, it still requires a study over a six-year period. This is a long time to track data and a dynamic university evolves during this time period. Therefore, we suggest that instead of using G_6 in the correlations studies suggested above, it might be both preferable and easier to define success by the percentage of students who have completed the first semester of university study in good standing. Also, we used the average ACT score of all students from any particular high school as a measure of high school quality. Other measures, like the percentage of high school students who go on to college, can also be tried if data is available.

References

Stern, D. and Briggs, D. (2001). Changing Admission Policies: Mounting Pressures, New Developments, Key Questions. *Change*, 33, 34–41.

*U.S. News and World Rep*ort Online (2001): http://www.usnews.com/usnews/edu/college/corank.htm

Tam, Mo Yin and Sukhatme, U. (2003). The Importance of High School Quality in University Admissions Decisions. *C and U Journal,* 3–8.

It is a pleasure to thank Carole Snow, Ann Feldman, Anne Cruz, Chris Messenger, Gilbert Bassett, Joseph Persky, Julie Smith and Barbara Zusman for helpful discussions and data.

Summary

With limited space and resources available at universities, college and universities must admit students with a good chance of being successful, especially for large, urban, public universities, where educational resources are funded in part by state tax revenue. We illustrated an easy approach to ascertain the importance of various input variables (student attributes) and incorporated many of them into the formation of a selection index. In our study, this approach motivated the use of improved variables as indicators of success. Among the variables we studied is HSPR*, a new variable which we specifically designed to take into consideration both the student's high school percentile rank and a measure of the quality of his/her high school. Our study showed HSPR* as a better indicator of future graduation than those attributes currently used in the admission procedure. Incorporating this variable into a selection index, with consideration of other criteria, will therefore result in better college admission decisions.

Footnotes

[1] One successful program at UIC is the SBC/Ameritech Scholarship Program assists minority undergraduate students in its College of Business Administration by providing them scholarships and mentors. In addition, the scholars are asked to be tutors for elementary students in the neighborhood surrounding UIC. Since its inception in 1999, the Program has assisted more than 240 students. Scholars in this program achieve higher than average GPA and significantly higher than average graduation rate.

[2] A least squares fit to the data is a simple regression process fitting a straight line (to the data), which minimizes the sum of the squares of the error of the fit. For this study, one straight line relates the six-year graduation rate to each of the various input variables.

[3] ACT score (ACT) lies in the range 0 to 36. We create a normalized student ACT score variable (NACT defined by NACT = ACT x 100 /36) that has values in the range of 0 to 100, just like the high school percentile rank.

[4] This is done by a least squares fit to a three-variable model where the six-year graduation rate is a function of two variables: HSPR and ACT.

[5] The normalized SI is composed by SI = p HSPR + (1-p) NACT, where $0<p<1$ and HSPR is the high school percentile rand and NACT is the normalized ACT.

[6] The normalized SI* is composed by SI = p HSPR* + (1-p) NACT, where $0<p<1$ and HSPR* is the modified high school percentile rand and NACT is the normalized ACT. Again, HSPR*=HSPR x NHSACT which takes into consideration the average high school ACT score.