Fall 2017

Moving from Forecast to Prediction: How Honors Programs Can Use Easily Accessible Predictive Analytics to Improve Enrollment Management

Joseph A. Cazier
Appalachian State University, josephcazier@gmail.com

Leslie Sargent Jones
Appalachian State University

Jennifer McGee
Appalachian State University

Mark Jacobs
Appalachian State University

Daniel Paprocki
Appalachian State University

See next page for additional authors

Authors
Joseph A. Cazier, Leslie Sargent Jones, Jennifer McGee, Mark Jacobs, Daniel Paprocki, and Rachel A. Sledge

Moving from Forecast to Prediction: How Honors Programs Can Use Easily Accessible Predictive Analytics to Improve Enrollment Management

Joseph A. Cazier, Leslie Sargent Jones, Jennifer McGee, Mark Jacobs, Daniel Paprocki, and Rachel A. Sledge
Appalachian State University

INTRODUCTION

Most enrollment management systems today use historical data to build rough forecasts of what percentage of students will likely accept an offer of enrollment based on historical acceptance rates. While this aggregate forecast method has its uses, we propose that building an enrollment model based on predicting an individual's likelihood of matriculation can be much more beneficial to an honors director than a historical aggregate forecast. Many complex predictive analytics techniques and specialized software can build such models, but here we show that a basic approach can also be easily accessible to honors directors where a small amount of data collection and

basic spreadsheet software allow them to capture most of the benefits without needing the skills of a data scientist.

The first step comes in understanding the difference between a forecast and a prediction. A forecast is an estimate of a future event, generally in aggregate form. For example, today I might forecast that our ice cream store will likely sell 1,000 scoops of ice cream based on weather, time of year, day of the week, and regional events—all useful information for staffing and inventory management as well as profitability analysis. Historically, an honors administrator might use this approach to predict the total number of students matriculating to the university or to an individual program.

However, with predictive analytics one can acquire even more detail that could be useful in a setting like an honors program where not just the total number of "customers" matter but which ones will create a well-rounded, diverse honors program with students from multiple backgrounds (Siegel).

In the ice cream case, a predictive analytics example might predict not just how many total ice cream scoops might be sold but how likely each individual is to buy ice cream. Deeper analysis might predict the type of ice cream, time of day customers might come, and how frequently they might visit the store. Predictive analytics might also lead to prescriptive analytics, where you learn what might be done to persuade someone who was not planning to buy ice cream to do so, e.g., what it might take to change a consumer's mind so that she will buy ice cream today or how we can we get her to buy two scoops instead of one or to bring a friend.

This type of predictive and prescriptive analytics has helped many organizations improve their efficiency and effectiveness (Siegel), and we believe that honors directors can also use it. In this approach, each potential honors student would receive an individualized probability score reflecting his or her likelihood of accepting an offer of admission. This score could still be aggregated into a direct forecast of how many students would likely attend, but it would also show the likelihood that any individual student would attend. The scores could predict how many from a certain group (e.g., science majors or Hispanic students) are likely to attend. This information could help strategically determine scholarship offers as well as the staff's time commitments to recruitment and follow-up activities.

BACKGROUND

An increasing amount of data is being collected about potential students when they apply for admission to a university. High school GPA, SAT scores,

and extracurricular activities are all part of the admissions application along with essays and other pieces of data about the student. Traditionally, this information has been used to determine if a student is prepared for university or qualified to attend (Mariz). However, developments in research methodology and technical capabilities have also made this information valuable in recruiting efforts and in extending the optimal number of enrollment offers.

University programs that depend on enrollment and completion can harness data analytics to manage yield and predict matriculation rates, significantly improving efficient use of resources. Like most programs, the honors program must sustain itself by predicting enrollment, attendance, and completion. While honors directors can examine applicant information to predict the best pool of potential students, the task may seem too daunting and time-consuming given their numerous other responsibilities. In addition to teaching a class or two, many honors directors manage entire programs and act as liaisons to others. Also, most honors directors are tasked with optimizing scholarship and resource usage. Directing scholarship awards to the students most likely to attend, while at the same time leveraging scholarship offers to attract the most qualified students, enhances an honors program's ability to enroll the most sought-after talent.

Factors Affecting Enrollment Decisions

Predicting an individual's overall likelihood of accepting an enrollment offer is precarious. Students who have not committed to a university usually have an idea of where they would like to go, but a scholarship offer can persuade them to enroll elsewhere. A 2013 study conducted by the University of California at Los Angeles concluded that financial aid offers affected the attendance decisions of 46 percent of the incoming freshmen, with 43 percent citing the overall cost of attendance as the main factor in their decisions (Pryor et al.). Honors directors thus play a pivotal role in attracting top students through scholarships and financial aid. Knowing a student's likelihood to accept an offer of admission to a given honors program may impact the strategic use of financial aid to build the best incoming honors class.

Using Analytics to Predict Enrollment

Historically, honors directors have relied on personal interviews, professional references, and written statements of intent in determining the likelihood of a student's accepting an offer. However, current capabilities

within the field of analytics allow more informed decisions about potential acceptance. Diverse data sets describing the high schools and geographic locations of subject students can connect student demographics with socio-cultural demographics. Matching these independent pieces of data to the characteristics of the university allows for a more granular examination of who is likely to attend and why.

Examining the characteristics unique to a particular honors program can also allow deepened predictive capabilities in the admissions process. For example, honors students tend to be goal-oriented academic achievers with specific reasons for choosing a university. Using a combination of each discrete piece of data, honors directors can estimate the probability of a student's accepting an offer, thus optimizing enrollment by improving yield management and recruiting efforts.

The Example Study

To illustrate the value of predictive analytics, our study uses data from a mid-sized regional university in the southeastern United States. At the time of this study, the university housed an honors program of approximately 800 students, with 150 entering freshmen and 80 internally recruited freshmen and sophomores in later semesters. The average student who completes the honors program has over a 3.45 GPA.

We use this example to describe the basic statistical and analytical methods employed in an analysis of factors that influenced accepted honors students' decisions to enroll. Other universities, honors directors, or similar programs can use the same basic process to predict more effectively enrollment rates among accepted students. The findings in this study show the importance of data integration in university recruitment and financial aid operations as well as the applicability of one university's methods to other institutions.

THEORETICAL BACKGROUND

Enrollment Management

Universities and honors programs struggle with extending enrollment and scholarship offers to students who are unlikely to attend based on a multitude of factors, including acceptance from a first-choice school, financial aid availability, and a student's preferred major (Pryor et al.). Enrollment management is the strategy used by universities and other institutions in estimating

an optimal offer pool that efficiently distributes financial aid opportunities and deploys effective recruitment efforts; it is an institutional response to the challenges and opportunities that recruiting and retaining the best student body composition presents to a university's financial health, reputation, and student quality (Baker).

Along with maximizing the academic profile of incoming student bodies, directors of enrollment management set goals such as increasing the population of the university, striving to diversify the university based on factors like race and socioeconomic status, and assuring availability of adequate housing and student affairs resources for incoming students (Martin & Moore). In pre-enrollment periods, the two main facets of enrollment management are recruitment and yield management. Recruitment encompasses a university's effectiveness in attracting desired students, and yield management describes the process by which enrollment directors optimize offer pools.

Recruitment Efforts

Honors and enrollment management directors must improve their recruiting efforts in order to attract the best students to their universities. As students are increasingly applying to schools that are farther away geographically, directors must take care about where they place recruitment resources ("Trends in Higher Education"). Some large universities with sufficient funding resources expend more resources for recruiting nonresident students in order to find potential high-quality freshmen and students who can afford to pay out of pocket (Jaquette & Curs). The mad dash for nationwide and worldwide recruitment creates intense competition among universities. Aggressive recruitment efforts through online and social media advertising, large financial-aid and scholarship packages, and value proposals based on a school's ranking are vital to capturing the greatest number of ideal students from across the globe (Burd).

Moreover, students in the United States are showing a declining perception of the value of a college degree ("Trends in Higher Education"). As potential university students become more skeptical about the value of a college degree, honors directors must become more creative in recruitment methods to regain potentially lost students. More precise and effective recruitment efforts also provide value to honors directors through cost reductions and resource optimization.

Given the importance of effective recruiting, honors directors should directly interact with students to communicate the competitive advantages

that their university enjoys over others (Ross & Carnes). As universities compete to differentiate themselves to potential students, recruiters must be dynamic in how they attract their desired candidate pools. Students' decisions on a choice of college are most likely influenced by campus-sponsored individual tours of universities along with the availability of early application benefits (Fischbach). Knowing the probability that a particular student or demographic of students will accept an enrollment offer is essential in deploying these recruitment techniques and can be improved by leveraging the data provided by potential candidates.

Yield Management

For honors directors, one of the greatest challenges is choosing the best offer pool with respect to scholarship and housing availability, demand projections, and the desired number of incoming students (Netessine & Shumsky). The precision with which honors programs make admissions offers is crucial in determining the success of the next class of students. Extending too many offers results in resources not being sufficiently distributed among the students and decreasing the value of the education the university provides (Green). In contrast, offering admission to too few students harms the competitive nature of the honors program, which should provide a close-knit community of students who perform better academically than their counterparts and go on to receive valuable research opportunities and internships (Cosgrove). Honors programs also incentivize students to achieve better grades and to participate in more challenging classes and extracurricular activities than their peers.

Example Student

To illustrate the process of predicting the probability that a student will enroll after an offer of admission, we will describe the individual characteristics of a sample student. The values for each characteristic are random. The sample student will have the following characteristics:

- HSGPA: 4.2
- SAT Score: 1380
- Intended Major: Accounting
- Residency Location: 1 (from the region in which the university is located)

- Gender: Female
- Race: African American
- Socioeconomic Status: From school with 30% free or reduced lunches.

Academic Credentials

This honors program in this case study determined an academic performance threshold that seemed reasonable and contained most of the students who had previously accepted enrollment offers from the institution. The academic performance threshold is simply a way to define the high school academic performance of incoming students. Students who held academic credentials higher than this threshold most likely would choose to attend different types of institutions, perhaps with higher academic requirements for acceptance. The threshold used in this case study was derived by conducting simple analysis on the distributions of GPA and SAT scores from students who accepted enrollment offers. The purpose of this stage in the case study was to provide the honors program with a target audience that would be most responsive to scholarship and enrollment opportunities. The program would also need to recruit more students above the threshold in order for the target number of students to attend.

GPA

Grade point average has historically been highly predictive of a student's performance in higher education and is an important component of all admissions criteria, especially for honors admissions. Not only do high school GPAs provide insight into students' capabilities, but they also indicate the amount of effort that students apply to their studies (Belfield & Crosta).

The honors program in this case study conducted an analysis of the range of GPAs among applicants who received enrollment offers. First, we had to partition the data into ranges, or bins, for analysis, and we tried to pick ranges that would have enough students to be significant but would be small enough to provide predictive power and granularity. Bin ranges were created with the primary goal of maintaining range uniformity and the secondary goal of having a similar but significant number of students in each bin.

Bin ranges are important because they divide the data into describable categories that contain information about a certain subsection of the data. For honors directors to determine their optimal bin ranges, they first find the distribution of GPAs among their accepted applicants. Some universities

experience a bell-curve type of distribution in which most of the applicants' GPAs are near the average while others observe skewed distributions with high or low GPAs.

Directors can create bins that contain uniform or cut-off ranges of GPAs while maintaining a similar number of students in each. Sorting the bin ranges in ascending order in a spreadsheet allows directors to easily determine the acceptance rates of each range. After sorting the spreadsheet so that each range is in the correct order, directors can find the acceptance probabilities for students in each range by averaging the students' acceptance responses. With 0 denoting a student who did not accept an enrollment offer and 1 signifying a student who did, averaging the series of 0's and 1's gives honors directors the acceptance probabilities of each range, which is simply the proportion of students in each sub-category.

Table 1 shows the results of the bin range analysis conducted by the subject honors program. As the applicant pools for each honors program are different, honors directors who conduct similar analyses may experience different results For this honors program, the first two bin ranges are extended 0.19 points to capture enough applicants in each. However, bins 3–10 each contain a uniform GPA range of 0.09 points. Column 3 indicates the number of students who earned a GPA within each range, and the fourth column contains the percentage of students within the range who accepted enrollment offers from the honors program. Percentages in Column 4 represent the

Table 1. Subject Honors Program's GPA Ranges and the Probabilities of Offer Acceptance Associated with Each

Bin	GPA	Offers	Acceptance Probability
1	3.81–4.0	38	50.0%
2	4.01–4.2	39	28.2%
3	4.21–4.3	36	50.0%
4	4.31–4.4	37	21.6%
5	4.41–4.5	56	32.1%
6	4.51–4.6	63	21.0%
7	4.61–4.7	69	21.7%
8	4.71–4.8	55	32.7%
9	4.81–4.9	54	20.4%
10	4.91–5.0	28	32.1%

probability that a student within that GPA range will enroll in the university. As GPA is not the only factor affecting predicted acceptance, these probabilities will be combined with probabilities derived from other factors to find cumulative individual probabilities. Honors directors can replicate the table above to summarize their results.

For our example student, we can calculate the probability of acceptance given her GPA of 4.2. Table 1 indicates the probability of her acceptance as 28.2% (bin 2). The bin from which this probability is derived is highlighted in Table 1 above, as will be true for the example student in all the Tables.

SAT Score

Compared to GPAs, some see SAT scores as a more direct measurement of an incoming freshmen's academic ability (Hannon & McNaughton-Cassill). Like high school students with high GPAs, students who earn high SAT scores are more likely to attend more competitive schools, leaving the middle- and lower-tier schools to compete aggressively for students with high SAT scores who do not attend upper-tier universities (Camara & Echternacht). Most universities can improve their recruitment techniques and resources for future incoming classes by knowing the probabilities of acceptance by students who earn certain SAT scores.

The process of creating the analysis for SAT scores and ranges is identical to the technique used for GPAs. First, honors directors can develop their own bin ranges based on the distribution of their data. Then, averaging the series of 0's and 1's in each bin range gives the probabilities of acceptance by students who earn SAT scores within those ranges. Table 2 shows the bins and the probabilities found by the subject honors program. Again, this table is merely a representation of the data that can be found by other programs, so the results will vary for each institution. The bin ranges were chosen to capture the best ranges with a similar number of observations in each. In the final analysis of individual probabilities, the percentages in this table will be combined with the GPA probabilities along with the percentages in the demographic variables.

For our example student, we can calculate the probability of acceptance given her SAT score. She had an SAT score of 1380, so the probability of her acceptance is 18.1% (bin 5).

Demographic Information

Gender

The first demographic variable that honors directors can leverage is gender. While there is little informative potential in studying gender distributions, directors can optimize their offer pools based on how many students of each gender they wish to attract in each class. As female students are the most numerous on most college campuses, including the one in this study, honors directors can adjust their populations by determining the probabilities of acceptance for each gender. According to the subject university's diversity website, 56% of the student population was female during the fall 2016 semester. Table 3 shows the genders and their probabilities for the subject honors program.

Honors directors can divide the dataset into male and female categories to average the acceptances. With 0 meaning a student did not accept and 1 meaning a student did, the average of the 0's and 1's gives the probabilities listed in the table. This university sees that females are more likely to accept

Table 2. Subject Honors Program's SAT Score Ranges and the Probabilities of Offer Acceptance Associated with Each

Bin	SAT Score	Offers	Acceptance Probability
1	1150–1300	61	49.2%
2	1300–1330	42	23.8%
3	1330–1350	83	32.5%
4	1350–1370	57	26.3%
5	1370–1390	83	18.1%
6	1390–1420	83	32.5%
7	1420–1460	67	25.4%
8	1460–1600	42	16.7%

Note: For Clarity, ACT Scores Have Been Converted to SAT Equivalents.

Table 3. Gender Distribution and Acceptance Probabilities of Each Gender at the Example Honors Program

Gender	Offers	Acceptance Probability
Female	332	29.2%
Male	186	27.4%

enrollment offers, but other programs may experience different results. The individual probabilities of male and female students will be combined with their academic credential probabilities along with the rest of the demographic variable probabilities.

We can calculate the probability of the example student's acceptance based on gender as 29.2%.

Race

Studying applicant pool race distributions provides a similar value as gender distributions. As campus diversity and inclusivity continue to emerge as important issues in college programs, directors can use acceptance probabilities to reach target populations (Hurtado). The example university's racial distribution is heavily concentrated in White students, who made up 84% of the student population in the fall 2016 semester. The next most represented racial group was Hispanics, who represented 4% of the student population. Table 4 shows racial distribution represented in the honors applicant pool and the probabilities that each would accept an enrollment offer. In the final cumulative probabilities, race will be included among the demographic variables, which will be weighed with the probabilities derived from the academic credentials.

Our example student is African American, so the probability that she will accept enrollment is 33.3%.

Major

The intended majors of honors applicants can also provide valuable information in predicting individual acceptance rates. As some universities

Table 4. Racial Distributions in the Studied Honors Program and the Offer Acceptance Rates of Each Race

Race	Race Code	Offers	Acceptance Probability
White	0	455	28.8%
African American	1	9	33.3%
Asian	2	14	7.1%
American Indian	3	13	23.1%
Hispanic/Mixed	4	18	27.7%
Hawaiian/Pacific Islander	5	1	0.0%

are renowned for particular programs of study, the probability that a student will attend a university based on major is a vital piece of information. Renowned programs are a university's points of pride, and directors recognize their prominence in creating higher acceptance rates. The university in this case study has a number of points of pride that anecdotal data suggest attract students:

- **Accounting**: This university's accounting students consistently boast some of the highest CPA pass rates in the nation. Accounting firms in the southeastern region of the United States aggressively recruit students from this school.
- **Anthropology**: This university maintains the largest undergraduate-only anthropology program in the state, and it is ranked third among the state's anthropology programs.
- **Economics**: The economics program is globally ranked in experimental and environmental economics. Economics students consistently win regional and national tournaments on economics topics.
- **Sustainability**: As the university is located in a mountainous region that affords opportunities to observe the natural environment, many students attend for the sustainability programs, which are nationally ranked on affordability and value measures.
- **Geology**: Like the environmental science program, the geology program attracts students who appreciate environmental education and outdoor activities. Geology majors are highly competitive in obtaining employment after graduation.
- **Math Education**: Founded as a normal/teacher's school, the university has a long reputation of maintaining a premier education program in the region. This program produces the most high school math teachers in the state.
- **Music Education**: The music program contains the state's Band of Distinction, an award it has earned consistently over the past few decades. Graduates from this program enjoy almost 100 percent job placement.

Determining points of pride is achievable through a good understanding of a university's strengths and weaknesses. For an honors director new to a

particular school, looking at the university's promotional materials and talking to faculty, parents, and prospective students might help with this process.

There may also be majors or programs that could make a student less likely to enroll. For example, a performing major that does not offer what students are looking for decreases the likelihood of their attendance. The example university does not offer an engineering degree, so students who want to be engineers are unlikely to attend. Likewise, an intended major in marine science is a negative indicator given the distance to the ocean and lack of a course of study in that area. Undecided majors might tell us something else. Again, the historical data can give a probability for each group of intended majors.

Table 5 shows a few of the university's departments, the number of students in each, and the probabilities of acceptance for each. The points of pride

Table 5. Subject University's Points of Pride or Weakness, Academic Departments, and Acceptance Probabilities for Each

A few Specific Areas	Broader Departments Taxonomy	Offers	Acceptance Probability
Accounting		7	14.3%
Anthropology		9	44.4%
Engineering		11	0.0%
Sustainability		16	31.3%
Geology		4	25.0%
Math, Secondary Education		7	71.4%
Music Education		14	50.0%
	Undecided	50	20.0%
	Education	25	32.0%
	Business	32	25.0%
	Theoretical Sciences	117	24.8%
	Applied Sciences	33	36.4%
	Fine Arts	26	38.5%
	Applied Arts	9	44.4%
	Environmental Science	11	9.1%
	Outdoors	23	30.4%
	Humanities	55	30.9%
	Behavioral Sciences	63	30.2%

and weaknesses are listed separately from the department codes to analyze acceptance probabilities based on what the university is known for. The numbers of students in the majors are then subtracted from the total number of students in all the departments; for instance, the business college (department code = 2) had a total of 39 applicants, but seven of them intended to become accounting majors so the number of intended business majors listed is 32. After arranging these areas into subsets unique to each school, honors directors can determine acceptance probabilities by averaging the 0's (students who did not accept) and 1's (students who did accept). These probabilities will be combined with the probabilities of the other variables.

In probability estimates, however, caution should be exercised to make sure the sample is a large enough in each category to make a generalized statement, perhaps by including data from previous years. The major alignment is one of the areas that has the most variability in matriculation rates and thus has the potential to be a significant indicator if the variability holds up over time.

The example student has indicated accounting as her desired major, so her probability of accepting enrollment is 14.3%.

Location

Location preferences and residency information are also important factors in a student's decision to accept an enrollment offer. Since the example university is located in a mountainous region, its candidates are likely to be interested in this environment. The university's reputation as a strongly performing regional university also attracts students from other states and countries. However, students from out of the state or country are among the least likely to attend.

While schools vary, this university's honors director found eight location distinctions, listed in Table 6. One distinction, region code = 0, is out-of-state and international students, who had similar matriculation rates in our sample. An honors director can look at various regions to see if they have different matriculation rates, aggregating those that are similar. This information can help determine where to spend scarce recruiting resources. Honors directors can determine region codes based on the size of the states in which they operate, the proximity of their location to other states and universities, and the populations of their states and regions. After determining the best region codes, finding the probabilities that students from these regions will accept enrollment offers is a matter of averaging the corresponding 0's and 1's.

We can calculate the probability of the example student's acceptance given her residency location. She is from Region 1, the region in which the university is located, and so her probability of accepting enrollment is 50%.

Socioeconomic Status

Students' socioeconomic status also provides information about their potential behaviors. The free and reduced lunch rates of the high schools that candidates attended are public data and can be a proxy for socioeconomic status. This information is generalized to a school's attendance zone, but it provides some potential cohort effects and information about life expectations.

In cohort effects, potential candidates are influenced by those peers who are not as likely to attend (Ransdell). This information provides an honors program with the likelihood that students from an area of low socioeconomic status will attend the university, and it signals a potential future need for scholarship opportunities and financial aid. Honors directors can use this information to leverage their financial resources in attracting desired students from underprivileged or affluent areas.

To determine the free/reduced lunch percentage (FRLP), the example honors program downloaded the state's free and reduced meals application data from the State Board of Education website which can be done in any state through a quick browser search for "free/reduced lunch applications in (insert state here)." The percentages found for each high school can then be matched with the students who attended them.

Table 6. Observed Regions from which Students Apply to the Subject University and the Acceptance Probabilities from Each

Region Code	Number	Probability
0	120	15.8%
1	18	50.0%
2	8	50.0%
3	28	42.9%
4	61	34.4%
5	150	21.3%
6	120	38.3%
7	13	38.5%

Table 7 shows the results from the subject honors program's analysis of free/reduced lunch percentage effects on acceptance probabilities. Bin ranges that are mostly uniform with similar numbers of observations were created for this analysis. The honors program received a large number of applicants from schools with 0% free/reduced lunch, forcing the creation of a bin that only contains students from those schools. The probability of offer acceptance seems to increase as the FRLP increases, indicating that students are more likely attend this program given its lower cost compared with other programs in the region.

We can calculate the probability of the example student's acceptance as 35.8% given the free/reduced lunch percentage at her high school of 30%.

THE COMBINED PREDICTION FORMULA

After determining optimal bin ranges and acceptance probabilities in each indicator, combined probabilities can be computed for each individual. The result is a spreadsheet that calculates the probability that a student or group of students will accept an offer from the program. Table 8 shows the cumulative probabilities for five applicants as an example. The SID column presents anonymous ID numbers applied to students to make the analysis possible. The GPA through Gender columns contain the acceptance probabilities that each student has based on the bin or category it is in.

For the sake of simplicity, we used an average of the probabilities indicated by each of the factors discussed earlier for the Cumulative Probability column. A more sophisticated analysis might assign weights to each variable. However, the goal was to provide some basic techniques to demonstrate their usefulness. For our example, the values in this column will provide the

Table 7. Free/Reduced Lunch Percentage Bin Ranges and the Offer Acceptance Probabilities of Each

Bin	FRLP	Number	Probability
1	0%	110	19.1%
2	1–10%	46	23.9%
3	11–20%	86	27.9%
4	21–30%	67	35.8%
5	31–40%	94	29.8%
6	41–50%	55	38.2%
7	51–80%	60	31.7%

probabilities that students will accept enrollment offers from the honors program. The "Accept?" column denotes whether the student actually accepted an enrollment offer. Honors directors can easily create this table in several ways. First, Excel VLOOKUP offers automated database input capabilities. With Tables 1 through 7 in separate spreadsheets or tabs in Excel, directors can use VLOOKUP to direct their software to place information from those tables into a table like Table 8. Using the SID numbers to provide a reference for the software, the process is fairly simple and quick.

Directors can also manually input the information from Tables 1 through 7 into a Table 8 format. This method is more time-intensive with copying and pasting but easy for smaller data sets if there is no interest in programming the VLOOKUP feature. Either way, converting the probabilities from each variable into cumulative probabilities is a process as simple as averaging a series of numbers.

Determining cumulative probabilities enables honors directors to optimize their enrollment pools. With a target number of students for an incoming class, an honors director can better identify the number of students who will receive enrollment offers to build the program that best serves the students and university goals. Table 8 allows directors to determine how many students are likely to accept offers based on the average probabilities. This number is derived by simply adding the cumulative probability columns together. Table 9 shows this process in action using the sample probabilities in Table 8.

As Table 9 shows, offering these five students enrollment into the honors program would yield 1.45 of them actually accepting. Obviously, the number of students who accept offers must be whole, but using this method gives honors directors a more detailed prediction of how many students will accept overall along with the specific probability of acceptance for each student.

Table 8. Cumulative Probabilities that Students Will Accept Enrollment Offers Based on Academic and Demographic Factors

SID	GPA	SAT	Location	Major	FRLP	Race	Gender	Cum. Prob.	Accept?
0001	0.282	0.181	0.158	0.302	0.191	0.288	0.292	0.242	0
0002	0.321	0.325	0.429	0.302	0.317	0.288	0.292	0.325	0
0003	0.282	0.181	0.385	0.500	0.358	0.288	0.274	0.324	1
0004	0.280	0.167	0.158	0.200	0.279	0.288	0.292	0.238	1
0005	0.210	0.325	0.383	0.364	0.382	0.288	0.292	0.321	1

With a complete dataset, an honors director can hone the model by updating the probabilities as each student accepts or rejects an offer. Directors can continue to offer enrollment to students until the total number is equal or approximately equal to their desired acceptance pool.

Furthermore, honors directors can repeat this process for each group so that they can have the probability, for instance, of all science students or a certain ethnic group and predict the total enrollment from each category. This approach makes it easier to manage diversity and program goals.

Example Student Cumulative Acceptance Probability

An example was given of each factor for calculating the acceptance probabilities of an example student. Table 10 outlines the individual probabilities for each factor and the cumulative probability that this student will accept enrollment. The cumulative acceptance probability is calculated by summing the probabilities for each factor and then taking the average.

The probability that this student will accept enrollment at the subject university is 0.324.

Table 9. The Total Number of Students Who Will Accept an Enrollment Offer Based on Their Cumulative Probabilities

SID	Cumulative Probability
0001	0.242
0002	0.325
0003	0.324
0004	0.238
0005	0.321
Total	**1.45**

Table 10. Combining the Factors for an Individual Prediction for Each Student

Sample Student	H. S. GPA	SAT Score	Gender	Race	Major	Loc.	FRLP	Predicted Cum. Prob.
Value	4.2	1380	Female	AA	Accounting	1	30%	-
Prob.	0.282	0.181	0.292	0.333	0.143	0.5	0.538	0.324

Examining Probabilities for Distinct Groups

Honors directors may want to increase the diversity of their student pool, which is a precarious balancing act if the diversity of many factors is addressed. Two examples might help illustrate the utility of predicting the number of offers required to ensure a certain number of students from various backgrounds.

Some honors directors may feel pressure to offer admissions to students intending to study a specific major. For example, if 10 honors students are required to enroll in accounting, the director could use the average acceptance rate for students intending to enroll in accounting to estimate how many offers they need to make. In subject university, the enrollment acceptance probability for a student intending to study accounting is 14.3%, meaning that the honors director would need to make approximately 70 offers to accounting majors to ensure that 10 would enroll.

The same method can ensure a healthy mixture of students from varying ethnic backgrounds. For example, an honors director may want to increase the number of Hispanic students. For the subject university, the acceptance probability of Hispanic/mixed students is 27.7%. Should the honors director need to increase the number of Hispanic honors students by 10, an additional 36 offers to Hispanic students would be required. In other words, if we multiply 36 offers of admission by the probability of enrollment, approximately 10 students will enroll.

CONCLUSION

A data-driven predictive approach can give honors directors information about which potential candidates might accept an offer of admission to their program, moving the process from a simple forecast to an individual prediction. Using this information can help honors directors make more informed decisions as they build their cohort.

While the method we have outlined is likely more sophisticated than many honors programs are using, it is not a perfect system. An ambitious honors director might take the analysis farther by using more predictive analysis, such as a decision tree or logistic regression algorithm to develop a segmented or weighted prediction that would balance the effect of each indicator variable and adjust for its importance. Once the data are collected, this approach could be completed relatively easily, but it is beyond the scope of

this paper. However, even our entry-level approach should improve on classical models of enrollment management.

Different schools have distinct factors that are important to them and their prospective students. The example provided is a guide for identifying those factors that can help predict what each individual applicant might do. Directors will want to update their model regularly as student populations change and new data come in.

While data analytics allow an honors director to hone acceptance procedures and present offers to applicants who are most likely to accept, they need adjustment for ethical considerations in order to create a fair and accurate procedure. An honors program should not turn qualified students away just because they do not match the typical matriculant. Diverse student populations do not always fit even the best statistical model. One way to avoid the ethical dilemma of making offers based on inherent demographic qualities would be to fill quotas based on the candidates most likely to accept an offer.

ACKNOWLEDGMENTS

Special thanks to our graduate assistants—LeAnne Byers, Julie Hayes, Tonia Olson, Sam Volstad, Stephen Kimel, Ben Sweeney, and Colin Cunningham—for their assistance on this project.

REFERENCES

Baker, C. H. (2012). Introducing Enrollment Management: Its Process and Structure. *The NAIS Enrollment Management Handbook: A Comprehensive Guide for Independent Schools* (pp. 5–9). Washington D.C.: National Association of Independent Schools.

Belfield, C. R., & Crosta, P. M. (2012). Predicting Success in College: The Importance of Placement Tests and High School Transcripts (CCRC Working Paper No. 42). New York, NY: Columbia University, Teachers College, Community College Research Center.

Burd, S. (2015). The Out-of-State Student Arms Race: How Public Universities use Merit Aid to Recruit Nonresident Students. Washington, D.C.: New America.

Camara, W. J., & Echternacht, G. (2000). The SAT[R] I and High School Grades: Utility in Predicting Success in College. Research notes. (College

Entrance Examination Board Rep. No. CB-RN-10). New York, NY: The College Board.

Cosgrove, J. (2004). The Impact of Honors Programs on Undergraduate Academic Performance, Retention, and Graduation. *Journal of the National Collegiate Honors Council, 5*(2), 45–54.

Elliott, K. M., & Healy, M. A. (2001). Key Factors Influencing Student Satisfaction Related to Recruitment and Retention. *Journal of Marketing for Higher Education, 10*(4), 1–11.

Epple, D., Romano, R., & Sieg, H. (2003). Peer Effects, Financial Aid and Selection of Students into Colleges and Universities: An Empirical Analysis. *Journal of Applied Econometrics, 18*(5), 501–25.

Fischbach, R. (2006). Assessing the Impact of University Open House Activities. *College Student Journal, 40*(1), 227–32.

Goodstein, L. (n.d.). The Challenge of Recruiting the Best. Retrieved from <http://gradschool.psu.edu/faculty-and-staff/practices/recruiting>

Green, K. C. (2011). The 2011 Inside Higher Ed Survey of College & University Admissions Directors. *Inside Higher Ed*. 1–32.

Hannon, B., & McNaughton-Cassill, M. (2011). SAT Performance: Understanding the Contributions of Cognitive/Learning and Social/Personality Factors. *Applied Cognitive Psychology, 25*(4), 528–35.

Hurtado, S. (2007). Linking Diversity with the Educational and Civic Missions of Higher Education. *The Review of Higher Education, 30*(2), 185–96.

Jaquette, O., & Curs, B. R. (2015). Creating the Out-of-State University: Do Public Universities Increase Nonresident Freshman Enrollment in Response to Declining State Appropriations? *Research in Higher Education*, *56*(6), 535–65. DOI: 10.1007/s11162-015-9362-2

Long, B. T. (2002). Attracting the Best: The Use of Honors Programs to Compete for Students. Harvard Graduate School of Education: unpublished working paper for the Spencer Foundation.

Maringe, F. (2006). University and Course Choice: Implications for Positioning, Recruitment and Marketing. *International Journal of Education Management*, *20*(6), 466–79.

Mariz, G. (2016). An Agenda for the Future of Research in Honors. *Journal of the National Collegiate Honors Council, 17*(1), 3–8.

Martin, J., & Moore, S. (2011). Strategies for Success within a Student Affairs-Based Enrollment Management Enterprise Custom: Research Brief. *The Advisory Board Company*. Retrieved from <http://www.uky.edu/ie/sites/www.uky.edu.ie/files/uploads/CF_Strategies%20for%20Success%20within%20a%20Student%20Affairs%>

Netessine, S., & Shumsky, R. (2002). Introduction to the Theory and Practice of Yield Management. *INFORMS Transactions on Education, 3*(1), 34–44.

Pryor, J. H., Eagan, K., Palucki, B. L., Hurtado, S., Berden, J., & Case, M. H. (2012). The American Freshman: National Norms Fall 2012. Los Angeles, CA: University of California, Los Angeles, Higher Education Research Institute.

Ransdell, S. (2011). There's Still No Free Lunch. *American Behavioral Scientist, 56*(7), 908–25.

Ross, J., & Carnes, L. (2012). Recruiting, retaining, and graduating college students. Retrieved from <http://business.eku.edu/sites/business.eku.edu/files/files/kbea/KBEA Journal Ross Carnes.pdf>

Siegel, E. (2013). Introduction. *Predictive Analytics: The Power to Predict Who Will Click*, 4–12. Wiley.

Titus, M. A. (2006). Understanding College Degree Completion of Students with Low Socioeconomic Status: The Influence of the Institutional Financial Context. *Research in Higher Education, 47*(4), 371–98.

Trends in Higher Education Marketing, Recruitment, and Technology. (2014, March). *Hanover Research*. Retrieved from <http://www.hanoverresearch.com/media/Trends-in-Higher-Education-Marketing-Recruitment-and-Technology-2.pdf>

The authors may be contacted at

josephcazier@gmail.com.