

# An Optimization Model for the Allocation of University Based Merit Aid

By Paul K. Sugrue

Paul K. Sugrue is Professor of Management Science at the University of Miami School of Business Administration in Florida

*The allocation of merit-based financial aid during the college admissions process presents postsecondary institutions with complex and financially expensive decisions. This article describes the application of linear programming as a decision tool in merit based financial aid decisions at a medium size private university. The objective defined for the model is to maximize the quality of the incoming class as measured by average combined SAT scores. The approach involves using the yield rates from the previous year for each combination of SAT score and merit aid award while eliminating from consideration combinations for which insufficient data is available. Parameter estimation is based upon 2006 data and the results of the model are measured against 2007 data. It is shown that the modeling approach yielded a higher average SAT scores when applied to actual 2007 admissions data.*

Student financial aid available in higher education can be divided into two distinct categories: university-based aid and external aid. External funds are all other sources of funds which are not university based, including funds from all federal, state, and other governmental agencies as well as any private sources. University-based aid derives from funds controlled solely by the university and may come from endowments or from operating budgets. It is not unusual for a private university to allocate close to one half of their gross tuition revenue to university-based financial aid. University aid funds can be further classified as designated and undesignated. As the name implies, designated aid is targeted for specific categories of students, while undesignated funds are directed toward the general student population. Undesignated university-based aid can be awarded to students based on their need, talent, or their academic merit. This article is concerned with the allocation of university-based undesignated aid awarded on the basis of academic merit only, which will be referred to as merit-based aid. Merit-based financial aid is also referred to as “non-need based” aid or more generally as “tuition discounting.”

On a macro level, the institution budgets a dollar amount for merit aid, whereas on a micro level, financial aid decision makers must allocate among accepted students the aid that has been budgeted. Before the allocation decisions are made, the objective in allocating the merit aid dollars must be articulated. Once the objective is determined, quantitative measures can be selected to measure the achievement of the objective. For example, if the objective is to improve the academic quality of the incoming class, then SAT score, class rank, or GPA can be selected as possible measures of achievement. These quantitative outcome measures give the decision makers at the micro level guidance for allocation decisions. Although a clearly stated objective and quantifiable outcome measures give guidance to the decision makers, the issue of how much merit aid to allocate to each student still remains.

The allocation of merit aid, that is how much merit aid to allocate to each student, is the topic of this article. This review assumes that the objective in allocating merit aid is to attract higher quality students in order to enhance the

overall academic stature of the incoming class and therefore the institution. Achievement of this objective is limited by the size of the class to be recruited, the availability of merit aid funds, and the availability of qualified applicants. For example, recruiting a class of 1,000 students by offering full tuition scholarships to all applicants with SAT scores above 1500 is most likely not feasible. In the first place, there probably would not be a sufficient number of applicants with SAT scores in that range. In the second place, there would most likely be insufficient funds to support this strategy, since the financial aid budget is normally derived from tuition revenue.

This article applies to the aid allocation decision the technique of constrained optimization. The approach is to formulate the problem as a mathematical programming problem that can be readily solved on a personal computer. Although the problem formulation is simple and straightforward, the challenge is in extracting the data required for the model and in interpreting the results. This study uses a subset of actual admissions data, which provides a clear example of how such a model could be successfully implemented at any university faced with merit aid allocation decisions.

## Background

The prediction of the yield rate, or the probability that an individual admitted student will enroll, is a major issue in the implementation of any decision model involving financial aid strategies. Earlier studies have examined the relationship between the yield rate and observable characteristics of applicants and of the institution for both a large selective university, as in Ehrenberg and Sherman (1984), and a smaller liberal arts college as in Moore, Studenmund, and Slobko (1991). In the examination of the elasticity of the fraction who accepted offers of admission at Cornell University, Ehrenberg (1984) concluded that financial aid plays an important role in the decision process, while Moore et al. (1991) estimated a positive relationship between the amount of the scholarship and the decision to enroll at Occidental College. While both studies examine the elasticity of the yield rate, neither provides guidance to the decision maker in improving the decision process. In fact, Ehrenberg (1984) cautions against the implementation of their results pending several years of observation of the required parameters, noting that the external environment facing universities is changing rapidly.

The application of mathematical programming models as decision making tools is pervasive in the field of business. In a now classic article, Robert Dorfman (1953) presented a clear non-algebraic exposition on the usefulness of mathematical programming in the solution of economic and business problems. The clarity of the examples in this article generated a slew of classic case studies in operation research courses. The use of linear programming as a decision tool in the financial aid decision was presented by Sugrue, Mehrotra, and Orehovic (2006). Specifically, the authors outline an approach to maximize the total net revenue while considering budget, student recruitment pools, SAT averages, and the enrollment targets as constraints.

The objective in making financial aid decisions varies with the type of aid awarded. Overall, the objective of the university may be to maximize net tuition revenue to the university as in Sugrue et al. (2006). In the case of need-based aid, the objective may be to minimize the financial strain on the families of the applicants. In the case of merit aid the objective is more generally associated with the overall quality of the incoming class. Need-based and merit aid awards are contrasted by Ness and Noland (2007), although their study

focuses on public funds. The objective of non-need based grants can be considered to enhance the enrollment management goals of the institution as described by Redd (2000).

In recent years there has been a move away from merit aid toward more need-based aid in some universities. Grossman (1995) described this trend among elite universities, making the argument that merit aid results in less money for needy students. However, one could argue that increasing academic stature is clearly not a major objective of the Ivies and Massachusetts Institute of Technology, which already sit atop the rankings of undergraduate universities. For middle-tier institutions, improving academic stature to rise in the rankings may be a major objective.

## The Linear Programming Model

The objective of the model described here is to maximize the average combined SAT score of enrolled students, while satisfying constraints on the availability of students within specified SAT ranges, the total amount of funds allocated to merit aid awards, and the target size of the incoming class. Linear Programming, a mathematical approach to solving constrained optimization problems, will be used to do so. The approach involves expressing the problem in terms of a set of linear functions which define the objective and the limits on the set of possible solutions. The first step is to define a set of variables whose values are to be decided. These variables are referred to as the decision variables. Once these variables are defined, the objective and each constraint can be expressed as a linear function of these decision variables. Solving a linear programming problem involves selecting values for the set of decision variables, within the set of feasible solutions, which either maximize or minimize the objective.

The decision variables for this problem formulation are the number of admitted students in a specific SAT range who are offered a specific merit aid award. Decision variable  $x_{ij}$  is defined as follows:

$x_{ij}$ : the number of students in SAT group  $j$  offered merit aid amount  $i$

If there are  $n$  ranges or groupings of SAT scores and  $m$  distinct merit aid award levels, then there would be  $(n*m)$  decision variables. Each possible solution to the problem will consist of the assignment of a value to each of the  $(n*m)$  decision variables.

The combination of an SAT range and a merit award level will result in a percentage or probability of applicants offered that combination enrolling. These probabilities vary with the SAT and award levels. For example, higher range SAT applicants would have a lower probability of enrolling for a given aid level than lower range applicant would. Equivalently, students in the same SAT range offered greater amounts of financial aid would have a higher probability of enrolling than those offered less aid. This probability of enrolling for a given combination of SAT range and merit award is  $y_{ij}$  and will be referred to as the yield rate. The yield rate  $y_{ij}$  is defined as follows:

$y_{ij}$ : the probability that applicants with SAT scores in range  $j$  and offered merit aid at level  $i$  will enroll

Unlike the decision variables  $x_{ij}$ , the values of  $y_{ij}$  must be known before the model can be solved. Such variables are called parameters of the model.

Each of the  $n$  groupings of SAT scores has a representative value (this could be the mean, midpoint, or median of the class). The representative value of each of the  $n$  SAT ranges is referred to as  $s_j$ , where  $s_j$  is defined as:

$s_j$ : the representative value of the SAT range  $j$

This representative value is computed from the SAT scores of all applicants who fall into the respective SAT range. These values would be known prior to merit awards decisions being made.

Each of the  $m$  merit award levels have a specific award level which will be referred to as  $a_i$ . Therefore the definition of  $a_i$  is as follows:

$a_i$ : the merit award level in dollars for merit group  $i$

Merit aid awards are typically given in set dollar amounts and therefore the number of groups is typically not large.

The expected value of the number of applicants who will accept the financial aid offer and enroll for a particular combination of SAT and award levels, is therefore  $(x_{ij} * y_{ij})$ , the number of students in SAT class  $j$  who are offered award  $i$  times the yield rate for that same combination. The expected value of the size of the class is the sum over all possible combinations of each of these individual expected values, or:

$$\sum_{i=1}^n \sum_{j=1}^m x_{ij} y_{ij}$$

If the target class size is  $C$ , the first constraint on values which can be assigned to the decision variables is:

$$\text{Constraint 1: } \sum_{i=1}^n \sum_{j=1}^m x_{ij} y_{ij} = C$$

This constraint states that the expected value of the class size must equal a predetermined value  $C$ .

There is almost always a limit on the total funds that can be disbursed for financial aid awards. In this model it is assumed that the total budget for merit aid awards is  $B$ . The expected expense of each award amount,  $i$ , for a given SAT level  $j$ , is the number of offers extended in that amount times the probability of student applicant in the given SAT range accepting the award times the amount of the award  $b_i$ . Therefore the total expected amount spent on all merit aid is the sum of the expected values of each award category. This total expectation can be expressed as:

$$\sum_{j=1}^m \sum_{i=1}^n b_i y_{ij} x_{ij}$$

If this total expected expenditure on merit aid cannot exceed the budget for merit awards, B, then the second constraint on values which can be assigned to the decision variables is:

$$\text{Constraint 2: } \sum_{j=1}^m \sum_{i=1}^n b_i y_{ij} x_{ij} \leq B$$

This constraint states that the expected value of all merit award offers made cannot exceed the merit award budget, B.

Prior to merit awards being extended, the number of applicants who have been accepted in each SAT range is known and therefore the number of offers extended in each range cannot exceed this number of accepted applicants. The number of accepted applicants in each range is  $p_j$  which represents the pool of candidates available for merit aid awards in SAT range j. The total offers made in each award category to students with a given SAT range cannot exceed the number of applicant accepted in that group. This can be expressed as:

$$\text{Constraints 3 thru } 3+n: \sum_{i=1}^n x_{ij} \leq p_j$$

One of these pool constraints would be required for each of the  $n$  SAT groupings.

In this model, the objective in assigning values to the decision variables will be to maximize the mean SAT score for the enrolled class. The mean of the enrolled class can be approximated by using a weighted mean of the SAT ranges using the expected number of students to enroll in each SAT range. If the representative SAT value for range j is  $s_j$  then the partial weighted mean of the SAT for range j can be expressed as:

$$\sum_{i=1}^n s_j y_{ij} x_{ij}$$

Summing the partial weighted means over all SAT groups and dividing by the sum of the weights, which is the size of the class C, gives an expression for the average SAT score for all enrolled students:

$$\text{Objective Function: } z = (\sum_{j=1}^m \sum_{i=1}^n s_j y_{ij} x_{ij}) / C$$

The objective function is a linear function of the decision variables for a set value of the class size C. The complete statement of the linear programming model is to:

Maximize:

$$z = (\sum_{j=1}^m \sum_{i=1}^n s_j y_{ij} x_{ij}) / C$$

Subject to:

$$\sum_{i=1}^n \sum_{j=1}^m x_{ij} y_{ij} = C$$

$$\sum_{j=1}^m \sum_{i=1}^n b_i y_{ij} x_{ij} \leq B$$

$$\sum_{i=1}^n x_{ij} \leq p_j \text{ for all SAT ranges, } j$$

$$x_{ij} \geq 0 \text{ for all } i \text{ and } j$$

In this model the target class size,  $C$ , the merit aid budget,  $B$ , and the pool sizes of admitted students in the various SAT groupings,  $p_j$ 's, are all known with certainty at the time that values for the decision variables must be set. The  $(n \times m)$  array of yield rates for the combinations of SAT grouping and merit award level must be estimated.

## Appreciation of the Model

The model was applied to a subset of actual data for the 2007-2008 academic year. The subset was chosen from all accepted students by selecting only student applicants who received four distinct merit award levels. These four merit award levels constituted 62.4% of the merit aid awarded in 2007. The merit award levels are shown in Table 1.

**Table 1: Merit Aid Award Levels**

Award Level (i)	Award amount (b <sub>i</sub> )
1	\$0
2	\$11,000
3	\$16,000
4	\$24,000

The SAT scores of the accepted students were divided into six groupings and are shown in Table 2. This table also shows the number of students accepted in each SAT group, or the  $p_j$ 's.

**Table 2: Pool Sizes of Accepted Students by SAT scores**

SAT Group (j)	SAT range	Group Mean (s <sub>j</sub> )	Pool size (p <sub>j</sub> )
1	below 1100	1023.15	240
2	1101-1200	1161.76	602
3	1201-1300	1263.16	1,667
4	1301-1400	1352.38	1,835
5	1401-1500	1443.44	921
6	1501-1600	1532.36	157

The actual yield rates observed for each combination of SAT group and Award Level are shown in Table 3. The number of awards offered is also shown in this table. It should be noted that in general these yield rates show the relationship between Award Level and SAT score that one would expect, i.e. as student quality increases, yield rates decrease and as award levels increase, yield rates increase.

**Table 3: Yield Rates by SAT Group and Award Level for 2007**

SAT Group (j)	Award Level (b <sub>i</sub> )			
	1	2	3	4
1	.458 (240)	.182 (11)	(0)	.667 (3)
2	.272 (393)	.244 (209)	.111 (9)	0.000 (3)
3	.248 (880)	.251 (438)	.235 (349)	.222 (9)
4	.159 (441)	.252 (345)	.226 (665)	.318 (384)
5	.090 (144)	.237 (38)	.161 (193)	.190 (584)
6	.056 (18)	.000 (2)	.071 (14)	.115 (157)

The number of awards offered appear in parenthesis for each combination of SAT score and award level

The budget amount, B, was the actual amount expended on the enrolled students who were given the four award levels listed above, which was \$12,967,710. The enrollment target, C, was the actual number of students who enrolled at the same four award levels, which was 1,281.

The 4x6 array of decision variables for this problem appears below:

$$X = \begin{bmatrix} X_{11} & X_{12} & X_{13} & X_{14} & X_{15} & X_{16} \\ X_{21} & X_{22} & X_{23} & X_{24} & X_{25} & X_{26} \\ X_{31} & X_{32} & X_{33} & X_{34} & X_{35} & X_{36} \\ X_{41} & X_{42} & X_{43} & X_{44} & X_{45} & X_{46} \end{bmatrix}$$

For example  $x_{23}$  represents the number of admitted students with an SAT score in the range 1201-1300 and offered a merit award of \$11,000. In the 4x6 array of yield rates is shown below:

$$Y = \begin{bmatrix} y_{11} & y_{12} & y_{13} & y_{14} & y_{15} & y_{16} \\ y_{21} & y_{22} & y_{23} & y_{24} & y_{25} & y_{26} \\ y_{31} & y_{32} & y_{33} & y_{34} & y_{35} & y_{36} \\ y_{41} & y_{42} & y_{43} & y_{44} & y_{45} & y_{46} \end{bmatrix}$$

$Y_{23}$  represents the probability that a student in the SAT range 1201-1300 who is offered a merit award of \$11,000 will enroll.

The arrays below show the actual yield rates and the number of observations for 2007. These yields and offer numbers are also shown in Table 3.

$$\begin{bmatrix} .458 & .272 & .248 & .159 & .090 & .056 \\ .182 & .244 & .251 & .252 & .237 & .000 \\ .111 & .235 & .226 & .161 & .071 & \\ .667 & .000 & .222 & .318 & .190 & .115 \end{bmatrix} \begin{bmatrix} 240 & 393 & 880 & 441 & 144 & 18 \\ 11 & 209 & 438 & 345 & 38 & 2 \\ 0 & 9 & 349 & 665 & 193 & 14 \\ 3 & 3 & 9 & 384 & 584 & 157 \end{bmatrix}$$

Eliminating cells with low observation numbers (less than 100), which is equivalent to eliminating that choice from consideration, gives the following yield rate matrix:

$$Y = \begin{bmatrix} .458 & .272 & .248 & .149 & .090 & & \\ & .244 & .251 & .252 & & & \\ & & .235 & .226 & .161 & & \\ & & & .318 & .190 & .115 & \end{bmatrix}$$

Now that all the parameters have been defined, the linear problem can be stated as:

Maximize:

$$Z = .37x_{11} + .25x_{12} + .25x_{13} + .17x_{14} + .10x_{15} + .22x_{22} + .25x_{23} + .27x_{24} + .23x_{33} + .24x_{34} + .18x_{35} + .34x_{44} + .21x_{45} + .14x_{46} \quad *$$

Subject to:

$$.458x_{11} + .272x_{12} + .248x_{13} + .159x_{14} + .090x_{15} + .244x_{22} + .251x_{23} + .252x_{24} + .235x_{33} + .226x_{34} + .161x_{35} + .318x_{44} + .190x_{45} + .115x_{46} = 1,281$$

$$2684x_{22} + 2761x_{23} + 2772x_{24} + 3760x_{33} + 3616x_{34} + 2576x_{35} + 7632x_{44} + 4560x_{45} + 2760x_{46} \leq 12,960,000$$

$$x_{11} \leq 240$$

$$x_{12} + x_{22} \leq 602$$

$$x_{13} + x_{23} + x_{33} \leq 1667$$

$$x_{14} + x_{24} + x_{34} + x_{44} \leq 1835$$

$$x_{15} + x_{35} + x_{45} \leq 921$$

$$x_{46} \leq 157$$

$$x_{11}, x_{12}, x_{13}, x_{14}, x_{15}, x_{22}, x_{23}, x_{24}, x_{33}, x_{34}, x_{35}, x_{44}, x_{45}, x_{46} \geq 0$$

The array on the left below shows the actual offers made to the fall of 2007 and the resulting average SAT score using 2007 actual yield rates. The array on the right shows the solution of the linear programming problem defined above with the resulting average SAT score again using 2007 actual yield rates.

$$\begin{bmatrix} 240 & 393 & 880 & 441 & 144 & & & & & & \\ & 209 & 438 & 345 & & & & & & & \\ & & 349 & 665 & 193 & & & & & & \\ & & & 384 & 584 & 157 & & & & & \end{bmatrix} \begin{bmatrix} 9 & 602 & 1667 & 0 & 0 & & & & & & \\ & 0 & 0 & 1167 & & & & & & & \\ & & 0 & 0 & 0 & & & & & & \\ & & & 668 & 921 & 157 & & & & & \end{bmatrix}$$

Average SAT = 1285

Average SAT = 1313

The aid allocation decisions based upon the model resulted in a 28 point increase in SAT score. The model result incorporated the 2007 yield rates shown previously, while to the decision makers these yield rates were yet unknown. In a real sense the allocations made by the model were under conditions of perfect information. In actuality, at the time the allocation decisions are being made the yield rates for the class being admitted (in this case the 2007 class) are not known. Therefore the input parameters of the model should be limited to information which was variable at the time the decisions must be made. One option is to use the yield rates from the previous year. Shown below are the yield rates from 2006 for the same combinations of SAT ranges and merit award levels:

$$\begin{bmatrix} .331 & .267 & .222 & .153 & .080 & & \\ & .259 & .219 & .192 & & & \\ & & .361 & .257 & .183 & & \\ & & & .336 & .204 & .137 & \end{bmatrix}$$

Using these 2006 yield rates in the model gave the following optimal solution and average SAT score:

$$\begin{bmatrix} 132 & 602 & 1667 & 0 & 0 & & \\ & 0 & 0 & 0 & & & \\ & & 0 & 1283 & 921 & & \\ & & & 552 & 0 & 157 & \end{bmatrix}$$

Average SAT = 1296

The actual 2007 yield rates were applied to the above solution set to compute the average of 1296. The linear programming solution resulted in a solution which had a mean SAT score 11 points higher than obtained by the actual allocations.

Applying 2007 yield rates to a solution based upon 2006 data shows other advantages of the linear programming solution. Table 4 below summarizes the differences between the actual allocation and the LP solution. Although the LP solution resulted in 12 fewer students enrolled, it generated \$969,560 more in net tuition revenue.

**Table 4: Comparison of LP model Solution and Actual Allocations**

	Actual Allocation	LP Model Solution
Average SAT	1285	1297
Enrollment	1281	1269
Tuition Revenue	\$39,352,756	\$39,012,614
Merit Aid Expenditure	\$12,967,710	\$11,658,008
Net Revenue	\$26,385,046	\$27,354,606
Financial Impact of LP Model		\$969,560

## Discussion and Conclusion

The allocation of merit-based financial aid during the college admission process presents award makers with complex and financially expensive decisions. Without a clearly stated objective, the common approach is to spend until all allocated funds are gone and then to ask for more. Applying linear programming to this class of problems presents a simple, straightforward, and disciplined technique for efficiently making allocation decisions. The example presented in this article showed a potential impact of \$1 million in a process involving the allocation of \$13 million. Most businesses would not hesitate to adopt such a competitive advantage, particularly when the cost of implementation is negligible.

The challenge in the implementation of such a decision aid is not in the availability of data or access to computer technology, but rather in training decision makers to understand and trust the power of this technique. Although the approach of using linear programming is straightforward and effective, the approach is extremely sensitive to yield rate estimation and the estimation of yield rates is not as straight forward. Yield rate estimation is a fertile area for future research.

### References

- Dorfman, Robert. (1953) Mathematical, or "linear," programming: A non-mathematical exposition. *The American Economic Review*, 18 (5), 797-825.
- Ehrenberg, Ronald G. and Daniel R. Sherman. (1984). Optimal financial aid policies for a selective university. *The Journal of Human Resources*, 19(2), 202-230.
- Grossman, Herschel. (1995). Scholarships: Need or merit? *Cato Journal*, 14(3).
- Moore, R. L., A. H. Studenmund and T. Slobko. (1991). The effect of the financial aid package on the choice of a selective college. *Economics of Higher Education Review*, 10(4), 311-321.
- Ness, E. C. and Noland, B. E. (2007) Targeted merit aid: i \ Implications of the tennessee education lottery scholarship program, *NASFAA Journal of Student Financial Aid*, 37(1), 7-17.
- Redd, Kenneth E. (2000). Discounting Toward Disaster: Tuition Discounting, College Finances, and Enrollments of Low-Income Undergraduates. *New Agenda Series: USA Group Foundation* 3 (2), December.
- Sugrue, Paul, Anuj Mehrotra, and Paul Orehovec. Financial Aid Management: An Optimization Approach . *International Journal of Operations Research* 1 (3), 267-282.