

Can Statistical Modeling Increase Annual Fund Performance?

*Paper presented at the 3rd Annual CASE/AIR conference,
St. Louis, Missouri, April 3-4, 2000*

Copies of this paper may be found at
<http://parsons.umaryland.edu/~porter/papers/CASEAIR2000.PDF>

Stephen R. Porter, Ph.D.
Evaluation Researcher
School of Nursing, University of Maryland, Baltimore
655 West Lombard Street
Baltimore, MD 21201
Phone: (706) 316-8618
Fax: (706) 316-0719
Email: porter@son.umaryland.edu

Abstract

Annual funds face pressures to contact all alumni to maximize participation, but these efforts are costly. This paper uses a logistic regression model to predict likely donors amongst alumni from the College of Arts & Humanities at the University of Maryland, College Park. Alumni were grouped according to their predicted probability of donating and then solicited for contributions during the current year's Annual Fund drive. Donation rates between likely and unlikely donors were not statistically significant from one another. Possible reasons for this null result are discussed.

Can Statistical Modeling Increase Annual Fund Performance? An Experiment at the University of Maryland, College Park

Introduction

Alumni donations have always been an important source of revenue for postsecondary institutions. Historically alumni provide about 25% of the voluntary support these institutions receive (Leslie and Ramey, 1988). Alumni donations have become even more important in recent years because the alumni participation rate is increasingly viewed as a measure of alumni satisfaction with an institution. *U.S. News & World Reports*, for example, uses alumni participation rates as one factor in their annual college rankings.

For large institutions this emphasis on participation rates puts them at a disadvantage. Alumni participation rates are calculated using the total number of alumni in the denominator, meaning that all alumni must be contacted to maximize the rate. But for an institution with tens of thousands of alumni, contacting each and every one of them can be a daunting and costly task.

Statistical modeling has the potential to efficiently separate likely from unlikely donors and thus aid in efforts to increase participation rates. There is a substantial body of statistical research on alumni donor behavior, but the emphasis is on explaining donor behavior. Yet advancement personnel are focused on *prediction*, not explanation, in their daily work. Unfortunately there is little work in the area of predicting donations. Consultants, who are notoriously close-mouthed about their methods, are doing most of the work in this area.

This paper tests the success of statistical modeling in discriminating between likely and unlikely alumni donors during an Annual Fund drive at the University of Maryland, College Park. The purpose of the paper is threefold.

First, the predictive power of the statistical models used in academia is tested. Such explicit testing is important because of a misconception that these techniques are "scientific", and therefore they are somehow better than the standard advancement approach of classifying donors by attributes such as past donation behavior. Both statistical models and advancement personnel do the exact same thing: they look at attributes of alumni to determine which alumni are likely to give money. The difference is that advancement personnel rely on their experience to weight the different attributes, whereas statistical models mathematically calculate a set of "best" weights.

Second, the paper demonstrates that experimentation is useful to academic administrators and is not merely something to be relegated to scientists in their laboratories. Too often administrators focus on the short term, emphasizing "let's get this done right away," rather than stepping back and questioning whether there are better ways to accomplish a task. It is generally useful to test new techniques and methods before implementation, and certainly any change in how alumni are solicited should be tested before use.

Finally, the paper shows that institutional researchers and advancement personnel can work together on projects of this type, and more importantly, that high-priced consultants are generally not needed. There is no mystery in what consultants do, whether it is predicting undergraduate acceptances during the admissions process or predicting alumni donations: they simply estimate a statistical model and use the results to make a prediction. With point-and-click statistical software like SPSS widely available there is no need for most institutions to hire outsiders. The statistical skills of institutional research offices combined with the data and substantive knowledge of advancement offices can be a powerful combination that should not be ignored.

Literature review

My review of the literature did not yield any specific articles detailing the use of statistical modeling to predict donation behavior. Articles in this area tend to be of two types, either academic research into why alumni give money to institutions (e.g. Bruggink and Siddiqui, 1995; Okunade, 1993, 1996; Taylor and Martin, 1995) or magazine articles describing the use of these methods to predict donations in general terms (e.g. Barth, 1998; Melchiori, 1988; Wylie, 1999). Table 1 provides a summary of the academic research and the primary explanatory variables used.

Table 1. Focus of Alumni Donation Research

Explanatory variables	Articles
Institutional attributes such as institutional quality, financial resources and fund-raising expenditures	Baade and Sundberg (1996b), Harrison (1995), Harrison et al. (1995), Leslie and Ramey (1988)
College athletics	Baade and Sundberg (1996a), Grimes and Chressanthis (1994)
Class year and reunions	Willemain et al. (1994)
Individual characteristics such as demographics, income and Greek participation	Bruggink and Siddiqui (1995), Okunade (1993), Okunade (1996), Okunade and Berl (1997), Okunade et al. (1994), Olsen et al. (1989), Taylor and Martin (1995)

Although the articles listed in Table 1 do not focus on predicting alumni donation behavior, they are very useful in identifying good predictors of likely donors. The statistical model used below employs some of the variables described in these papers.

The experiment

After meetings with representatives from the Annual Fund, a plan was developed to test the ability of statistical modeling in identifying likely donors. The university grants approximately 5,000 bachelor degrees per year, which means the annual fund has over 150,000 alumni to contact every year. The Annual Fund segments these alumni into eleven groups based upon donation history, Alumni Association membership, etc. The largest group is the segment called "Non-Donors". These are alumni who have not given to the Annual Fund in the last four fiscal years. Historically the participation rate for all alumni is 16%, while by definition the most recent participation rate for Non-Donors is 0%. However, every year some of the alumni in the Non-Donor group do pledge money when contacted during the Annual Fund drive.

These alumni pose a problem for the university from a solicitation point of view. In order to maximize the participation rate, all of these alumni must be contacted. Yet given their large numbers, considerable resources are expended to contact a group of alumni who are very unlikely to participate in the Annual Fund drive. Thus members of this group are excellent candidates for use in statistical modeling. If the few donors in this group could somehow be identified, contacting all alumni in the group would be unnecessary. Instead, these solicitation resources could be shifted to more profitable areas.

The Annual Fund staff contact alumni by college, and they agreed to use the next upcoming college in the current Annual Fund drive, Arts and Humanities (A&H), as the basis for the experiment. The procedure used was as follows. A logistic regression model predicting whether or not an alumnus would give money to the university was estimated, and a predicted probability of giving was assigned to each alumnus in the Non-Donor segment. The alumni in

this segment were then divided into two groups, Low and High, based on the median of this probability.

These two groups were further subdivided by the number of times alumni would be contacted. Usually the Annual Fund attempts to contact alumni ten times. Varying the number of contacts by group would give some idea as to the effect of multiple phone calls. Because members of the high probability group would be more likely to give, we decided to test the impact of increasing the number of contacts to fifteen. And because the low probability group would be less likely to give, we decided to have a group that would only be contacted five times instead of the normal ten. The result was four groups of alumni:

- Low 5** - low probability, maximum of 5 calls
- Low 10** - low probability, maximum of 10 calls
- High 10** - high probability, maximum of 10 calls
- High 15** - high probability, maximum of 15 calls.

Members of the low and high probability group were randomly selected and placed into the five or ten and ten or fifteen contact groups. Each group was given a target contact rate of 70%, so that with the exception of the number of contacts they would undergo the same experimental conditions.

Based on the predicted probabilities and the number of contacts, we have three hypotheses to be tested about the donation rates for each group:

- H1:** High 10 > Low 10
- H2:** Low 10 > Low 5
- H3:** High 15 > High 10

Methodology

Some researchers use alumni survey samples for their datasets (Bruggink and Siddiqui, 1995; Okunade, 1993; Taylor and Martin, 1995). Yet most alumni surveys have substantial non-

response bias: the response rates are so low that the alumni who answer the survey are not representative of the entire population of alumni. For example, alumni who answered the Class of 1994 Alumni Survey at the University of Maryland were twice as likely to give money to the university than the entire class. Rather than rely on survey data, the data used here are taken from institutional student databases for all alumni.

Occasionally researchers truncate their datasets and only examine alumni who donate (Okunade et al., 1994; Okunade, 1996; Taylor and Martin, 1995). Such truncation produces biased statistical estimates (Greene, 1997), in turn resulting in poorer predictive ability. Common sense tells us why this is problematic: if we want to understand and predict donor behavior, we must be able to distinguish between donors and non-donors, not just between low and high donors. The data used here include both donors and non-donors.

A model based on variables used in previous research on alumni donations was developed using variables from institutional databases. Unfortunately College Park's databases only contain data on students since 1978. This meant that 55% of the A&H alumni could not be used in the statistical modeling phase due to a lack of student data. These alumni were also solicited during the Annual Fund drive and are referred to as “Unknown” in the paper.

The dependent variable used in the model was a binary variable indicating whether the alumni had ever donated money to the institution, either during an Annual Fund drive or during some other solicitation. In the college as a whole 5.5% of the alumni had given money to the university. The independent variables can be divided into two groups, the first attempting to measure school experiences and the second measuring circumstances after graduation.

Matriculation status, taking the value of one if the alumnus entered the university as a freshmen, zero if a transfer, proxies the attachment of the alumnus to the university. Transfer

students spend less time at the university than students who enter as freshmen, and thus may be less likely to develop a deep attachment to the university. In addition, an internal university study found that transfer students are in general poorly treated compared with freshmen, again implying differential feelings towards the institution. A dummy variable measuring whether the alumnus had ever belonged to the Alumni Association also proxies attachment to the institution (Taylor and Martin, 1995). Because alumni who live closer to the university may be more involved than those living farther away (Bruggink and Siddiqui, 1995) distance in miles between alumnus' residence and the university is included in the model.

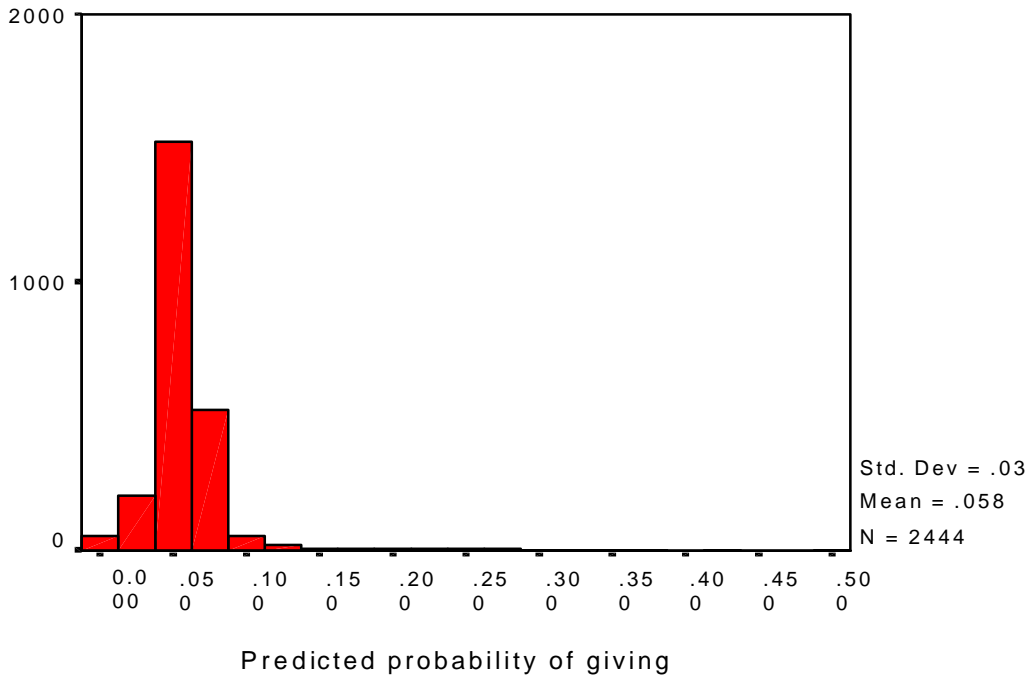
Four variables measure student success at the institution: cumulative grade point average, multiple majors at graduation, multiple degrees at graduation (both dummy variables), and time to degree, measured by the number of years between matriculation and graduation. Students who were successful at the university should be more likely to make a donation than students who performed poorly (Okunade et al., 1994).

The second set of variables attempts to measure the economic status of the alumnus. Income is not surprisingly a strong predictor of donor behavior (e.g. Okunade and Berl, 1997). The median income for the alumnus' zipcode of residence (based on 1993 Census Bureau projections) acts as a rough proxy for income. Two demographic variables proxy income also. An indicator variable for female alumni was included to control for differences in income between men and women (Baade and Sundberg, 1996a, 1996b).

One of the common findings in the alumni donation literature is a life cycle effect, where alumni are reluctant to give when young, give more as they age and accumulate wealth, and then reduce giving as they prepare for retirement (Olsen et al., 1989). To measure this effect the alumnus' age in years and age in years squared (to account for nonlinearity) are included.

group and 7.4% for the High group. In contrast, predicted probabilities for alumni in the entire college are as large as 75%.

Figure 1. Distribution of Predicted Probabilities for Non-Donors



Solicitation results

For several weeks in February and March alumni in the four groups were called and solicited to participate in this year's Annual Fund drive. Solicitation of each group was targeted to end when 70% of the group were contacted (members of each group were randomly contacted). The results of this solicitation effort are presented in Table 3.

The first three columns of the table describe the test groups and sample N's. The completion rate for each group is listed in the fourth column. Although each group was given a target completion rate of 70%, this goal was not met for the Low 5 and Unknown groups due to time constraints. As can be seen, there is not much variation amongst the groups for the

Table 3. A&H “Non-Donor” Target Groups and Solicitation Results

Groups			Outcomes				
Probability of donation	Mean number of calls	N	Completion rate	Number of donors	Percentage of donors	Mean donation	Median donation
High	10	611	72.5%	37	6.1	\$46	\$25
High	15	611	69.9%	26	4.3	\$42	\$25
Low	5	611	63.8%	24	3.9	\$49	\$25
Low	10	611	70.4%	38	6.2	\$39	\$25
Unknown	10	981	59.2%	40	4.1	\$34	\$25
TOTAL	10	3425	66.3%	165	4.8	\$41	\$25

percentage of alumni making a donation. The mean and median donation amounts are also very similar.

Statistical tests were performed to test the three hypotheses described above: the High 10 group would donate at a higher rate than the Low 10 group, the Low 10 group would donate at a higher rate than the Low 5 group, and the High 15 group would donate at a higher rate than the High 10 group. Table 4 shows the results. Without exception, the null hypothesis that the donation rates for the two groups being tested are equal cannot be rejected. The data indicate that the low probability and high probability groups have equal donation rates, and that varying the number of contacts did not have an effect on donation rates.

Table 4. Test of Differences between Donation Rates

Hypotheses	Donation rates	T statistic	P value
High 10 > Low 10	6.1, 6.2	0.12	0.905
Low 10 > Low 5	6.2, 3.9	1.83	0.068
High 15 > High 10	4.3, 6.1	1.42	0.155

Why the null result?

Most obviously, the model may not do a good job discriminating between likely and unlikely donors due to poor predictor variables. This is probably not the case, due to the fact that the variables used are common in the academic literature, several of them are statistically significant, and the model appears to do a good job predicting donations for the entire college.

More likely, it could be that the donor rates are so low for this segment that donor behavior is essentially random. Random in this sense means due to effects not included in the model and that would be difficult to measure and include. For example, an alumnus may have donated money simply because she received a raise that day. When donor rates are substantial this random behavior is merely noise, but when they are low donations due to such random behavior becomes a much larger proportion of the donor pool.

Future research

The results presented here would indicate that statistical modeling might not be a useful tool for predicting alumni donor behavior. But this conclusion is premature, given the segment of donors used in the sample. Traditional advancement techniques had also failed to identify likely donors in this segment of alumni.

More experimentation is needed before we can reject the use of these models. Perhaps the most glaring drawback of the model used here is the exclusion of time as a dimension (this was due to both data and time (no pun intended) constraints). Donor behavior varies over time, and including time in our models should significantly increase our predictive ability. Inclusion of time introduces methodological complications, but they are not surmountable. Examples of good work in this area are Bruggink and Siddiqui (1995) and Lankford and Wyckoff (1991). In

addition, time-varying variables should be included, such as the tax benefits to charitable donations (Feldstein and Taylor, 1976) and the state of the economy (Okunade, 1996).

The final lesson of this paper is the need for good data. Most institutions are probably similar to College Park in that they first computerized their student data in the early 70's. Data prior to this period are in microfilm form, and the benefits of having these data are not worth the costs of entering this data into computer systems. Of greater concern are the student data that have been entered. Some institutions may not be archiving these data because so much time has passed since the students involved attended the institution. Recently at College Park, for example, the data administrators floated the idea of discarding the early student computer data to save on storage costs.

Statistical models of donor behavior are useless without good data. Given that one of the strongest findings in the academic literature is the relationship between age and giving, and that we lack data for most of our middle-aged and older alumni, it is essential that institutions preserve their older student data.

References

- Baade, Robert A. and Jeffrey O. Sundberg. 1996a. Fourth Down and Gold to Go? Assessing the Link Between Athletics and Alumni Giving. *Social Science Quarterly* 774:789-803.
- Baade, Robert A. and Jeffrey O. Sundberg. 1996b. What Determines Alumni Generosity? *Economics of Education Review* 151:75-81.
- Barth, Steve. 1998. Finding the Needle in the Haystack: Use Computer Screening and Database Analysis to Discover the Hidden Major-Gift Prospects Among Your Alumni. *Currents* June: 32-38.
- Bruggink, Thomas H. and Kamran Siddiqui. 1995. An Econometric Model of Alumni Giving: A Case Study for a Liberal Arts College. *American Economist*, 392:53-60.
- Feldstein, Martin and Amy Taylor. 1976. The Income Tax and Charitable Contributions. *Econometrica* 44(6): 1201-1222.
- Greene, William H. 1997. *Econometric Analysis* New York: MacMillan Publishing Company.
- Grimes, Paul W. and George A. Chressanthis. 1994. Alumni Contributions to Academics: The Role of Intercollegiate Sports and NCAA Sanctions. *American Journal of Economics and Sociology* 531:27-40.
- Harrison, William B. 1995 College Relations and Fund-Raising Expenditures: Influencing the Probability of Alumni Giving to Higher Education. *Economics of Education Review* 141:73-84.
- Harrison, William B., Shannon K. Mitchell and Steven P. Peterson. 1995. Alumni Donations and Colleges' Development Expenditures: Does Spending Matter? *American Journal of Economics and Sociology* 544:397-412.

- Lankford, R. Hamilton and James H. Wyckoff. 1991. Modeling Charitable Giving Using a Box-Cox Standard Tobit Model. *Review of Economics and Statistics* 733: 460-470
- Leslie, Larry L. and Garey Ramey. 1988. Donor Behavior and Voluntary Support for Higher Education Institutions. *Journal of Higher Education* 592:115-132.
- Melchiori, Gerlinda S. 1988. Ranking Achievement: Use Your Alumni Data to Identify Your Best Prospects. *Currents* July/August: 52-56.
- Okunade, Albert A. 1993. Logistic Regression and Probability of Business School Alumni Donations: Micro-data Evidence. *Education Economics* 13:243-258.
- Okunade, Albert A. 1996. Graduate School Alumni Donations to Academic Funds: Micro-data Evidence. *American Journal of Economics and Sociology* 552:213-229.
- Okunade, Albert A., and Robert L. Berl. 1997. Determinants of Charitable Giving of Business School Alumni. *Research in Higher Education* 382:201-214.
- Okunade, Albert A., Phanindra V. Wunnava and Raymond Walsh, Jr. 1994. Charitable Giving of Alumni: Micro-data Evidence From a Large Public University. *American Journal of Economics and Sociology* 531:73-84.
- Olsen, Katherine, Amy L. Smith and Phanindra V. Wunnava. 1989. An Empirical Study of the Life-Cycle Hypothesis with Respect to Alumni Donations. *American Economist* 332:60-63.
- Taylor, Alton L. and Joseph C. Martin, Jr. 1995. Characteristics of Alumni Donors at a Research I Public University. *Research in Higher Education* 363:283-302.
- Willemain, Thomas R., Anil Goyal, Mark Van Deven and Inderpreet S Thukral. 1994. Alumni Giving: The Influences of Reunion, Class and Year. *Research in Higher Education* 352:201-214.

Wylie, Peter B. 1999. Model Behavior: Statistical Modeling Can Help You Find the Right Equation for Annual Fund Success. *Currents*, April.