A Logit Model with Interactions for Predicting Major Gift Donors

Share Your Story

Accessibility

A LOGIT MODEL WITH INTERACTIONS FOR PREDICTING MAJOR GIFT DONORS

Wesley E. Lindahl and Christopher Winship

....
....

We provide a new statistical model developed from the alumni database at North-western University for identifying potential major gift donors. Our logit model with interactions predicts which individuals will give $10,000 or more over three years using CHAID (Chi-square Automatic Interaction Detection) combined with logit analysis. Our work goes beyond our earlier research (Lindahl and Winship, 1992) by providing a more effective means of selecting fund-raising prospects. The critical new component in our model is the inclusion of interactions between past giving and other variables. Specifically, we find that for individuals with different past giving records different variables are important. Our results show that prospects with a low past giving level will rarely give gifts over $10,000.

....
....

With the increased importance of philanthropy to higher education and the increased demand by the public for efficiency in fund-raising, it is particularly valuable to investigate models that can serve to focus fund-raising efforts on those segments of the population that provide the greatest potential benefits at the lowest costs. Over the past three years we have developed models to predict donor giving at different levels—major gifts and annual gifts (Lindahl and Winship, 1992). These models include both internal data (past giving, peer review gift potential ratings, volunteer activity) and external data (questionnaire responses, geodemographic coding). Although the full models tend to be superior overall, a model including only past giving accounts for most of the predictive power. Our previous models used independent variables without allowing for interaction effects.

In our new research we develop models that incorporate interactions among the various independent variables. Specifically, we examine whether certain variables have predictive power for only particular groups of individuals. For example, does the salary level of a prospect only matter when the prospect is

Wesley E. Lindahl, Office of Development and Alumni Information Services, Northwestern University, 2020 Ridge Ave., Room 230, Evanston IL 60208-4305. Christopher Winship, Harvard University.

young? Is the affluence of the prospects' neighborhood only a valid predictor when the prospect has volunteered in the institution? These are both examples of interaction effects, in which the effect of one particular variable is affected by or "interacts" with another. In other words, the effect of one variable depends on the level of another variable. By including interaction effects, a model may better predict who will be a major gift donor.

We find that the interaction of one variable, past giving, with all other variables is of dominant importance. For high-level past givers (average giving = > $500), recency of past giving and size of past gifts are the most important factors. For moderate-level past givers ($499 = > average giving = > $100), age, capacity (salary and peer review gift potential ratings), and interest in planned giving are important. For the low-level past givers (average giving < $100), the same variables as in the moderate giving level model are important as well as other variables, including religious activity as a student and participation in university programs. Our results show that prospects with a low past giving level will rarely give gifts over $10,000.

OTHER INSTITUTIONAL AND EXTERNAL FACTORS

Two other possible independent variables were considered, but not included in the model. The first was development staff contacts. Unfortunately we lack good information regarding contacts over the past few years. Hopefully, we can rectify this in the future, so we can examine the importance of contacts for fund-raising success.

The second variable, an external factor, was the income tax rate. Weisbrod (1988, pp. 93–94) summarizes the evidence of the effect of changes in taxes, or equivalently the price of giving (the higher the tax rate the lower the price of giving):

> The mean price elasticity estimated from these studies is - 1.24; thus, the apparent average responsiveness of individual charitable giving to changes in income tax rates is relatively great, a given percentage change in the "price" of donating results in a 24 percent greater change in donations. . . . The wide range of elasticity estimates is a basis for some concern as to the appropriateness of the models and data used to generate them.

All the studies reviewed by Weisbrod disregarded the reputation and level of fund-raising expenditure of the recipient organizations. Motivational studies have shown that major donors place tax issues low on the list of influential factors. Panas (1984) found that tax incentives were often the least important consideration of major donors. Hunter (1975) interviewed 30 donors of $1 million or more, asking them to tell what motivated them to make such gifts. Their primary responses were worthiness of the cause, personal interest or association with the cause, knowledge that the organization was managed well, a sense of real social need, a sense of community obligation, and tax benefits.

Even considering the mixed evidence for the influence of the tax rates on donor giving, the variable would be helpful to include in donor giving models. However, the tax brackets of individual alumni were not available to us. Salary information in part provides a proxy measure of an individual's tax rate. Given that our interest is predicting which individuals are most likely to make a gift, the inclusion of tax rate in addition to salary would probably improve our model marginally. Recent changes in the tax law might change the likelihood that individuals will make gifts, but are unlikely to change the rank ordering of individuals on this dimension. Since the effect of changes in the tax rate to giving behavior takes hold slowly over the years (Vickrey, 1975, p. 157), the effect of not adjusting for changes in tax rates should be small. Furthermore, since the large adjustment in tax rates in the early 1980s, the rates for high-income donors have been relatively stable during the time period of the study. The new top tax rates recently passed by the United States Congress, although slightly higher than the current levels, are still fairly "flat" compared to pre-1981 rates.

PAST RESEARCH

Brittingham and Pezzullo (1990, pp. 39–44) provide an excellent summary of the predictors that have been considered by researchers over the past years. They divide the predictors into two major groups: characteristics of alumni when they were students and current characteristics of alumni. Only the latter have shown to be strong predictors. There was no conclusive pattern to support differences based on entering ability, patterns of attendance, participation in student organizations, place of residence, choice of major, grade-point average, or financial aid.

Brittingham and Pezzullo note that certain current characteristics of alumni were found to be predictors in some studies, but not others. Income, age, number of degrees from the institution, emotional attachment to the school, participation in alumni events, and participation in and donation to other voluntary and religious groups were found to be predictors. The areas where studies disagreed were the following: sex, marital status, spouse also alumnus, status of children, distance from school, and occupation. All of these characteristics were considered in our earlier research and those that did show significant and relatively strong effects were included for consideration in this study. Past giving, used as a predictor in our earlier research, was also included.

Most of the past research done predicting donor giving behavior makes use of linear techniques. Connolly and Blanchette (1986) used discriminant analysis, and Gerlinda Melchiori (1988) used classification analysis to predict donor behavior, both of which are types of linear regression. These techniques are inappropriate when the object is to predict rare events (such as giving over $10,000) or when the dependent variable has an upper or lower bound and

there are a large number of individuals at the bound (as with giving, where there are numerous individuals with zero giving). Logit analysis overcomes these problems. A full description of our past research and other research in this general area is available in our previous paper (Lindahl and Winship, 1992).

In the next section of the paper we describe our data, the variables we use, and the analysis techniques. We then present our results, first examining the results of our CHAID model and then exploring our logit estimates. We then test our model against other models by examining how well it does at actually predicting donors.

METHODOLOGY

The approach here is similar to that of our earlier research projects (Lindahl and Winship, 1992). We use the last three years of giving as the dependent variable and earlier giving along with internal and external variables to develop our models. Our data are analyzed in two steps. First, we use CHAID (Chisquare Automatic Interaction Detection) procedure, a technique that has been used in marketing research in the for-profit sector. CHAID partitions data into mutually exclusive subsets that best distinguish between groups along the dependent variable. This partition of the data indicates interaction effects that are likely to be needed in a model for predicting donors. The technique is described in detail by Kass (1980). In the second step we use logit analysis to actually predict giving, incorporating the interactions suggested by the CHAID analysis into our model. This step involves the development of a formal model for predicting donors.

Description of Database

We used the Northwestern University alumni/development database in this study. It consists of approximately 190,000 individuals with good current addresses, 140,000 of whom are alumni. Our analysis only uses the alumni. The information in the files has been assembled over the past twenty-five years from several sources. Questionnaires have been sent to alumni from undergraduate and professional schools at various intervals from the late seventies to the present. A few years ago, Northwestern carried out a National Resources Program that involved peer evaluation of the gift-giving potential of alumni; approximately 25,000 alumni were rated, some multiple times. Each year current graduates are added to the system, increasing the data base by approximately 4,000 records.

Geodemographic wealth rating information, provided by John Grenzebach and Associates from the Claritas Corporation, was appended to the main file in

1990. These variables are derived from census block level data. PRIZM (brand name of the Claritas product) clusters provide a market segmentation system that is based on the principle that people with similar backgrounds, means, and consumer behavior cluster in neighborhoods suited to their chosen lifestyles. Through a statistical analysis of the demographic characteristics and actual consumer behavior, every neighborhood in the nation is assigned to one of the 40 PRIZM clusters.

The three PRIZM clusters used in the logit model are labeled "Blue Blood Estates," "Urban Gold Coast," and "Grey Power." Claritas provides the following descriptions of these neighborhoods in its promotional literature. Blue Blood Estates are America's wealthiest socioeconomic neighborhoods, populated by super-upper-established managers, professionals, and heirs to "old money," accustomed to privilege and living in luxurious surroundings. Urban Gold Coast is the most densely populated cluster, with the highest concentration of one-person households in multiunit, high-rise buildings, and the lowest incidence of auto ownership. Grey Power comprises the nation's most affluent elderly, retired, and widowed neighborhoods, where nearly two million senior citizens who have chosen to pull up their roots are retired amongst their peers.

Analysis Design

We used total giving of $10,000 or more between 1988 and 1990 as the dependent variable, and giving prior to 1988 as an independent variable. We took the information that would have been available in December 1987, and used it to try to predict which alumni were most likely to give $10,000 or more between 1988 and 1990. For the nongiving variables, only currently stored values are known, so these are what we used. Since these variables (except for address information) have no dates associated with them, we assume that they have not changed in the past three years. Obviously, for some variables this is not problematic (for example, years since Northwestern degree) whereas for others changes may have occurred (for example, salary).

The use of the $10,000 level instead of the $100,000 level as employed in our earlier paper (Lindahl and Winship, 1992) stems from two factors. First, CHAID analysis using the $100,000 level encountered sample size problems since only 61 people in the file showed gifts over $100,000 in three years. At this level, CHAID showed a very large effect for past giving for those first few splits, which was predictable from our earlier research. Using the lower level of $10,000 allowed CHAID to do a more refined division of the data and reveal several key interactions.

Second, most institutions have the development staff available to work at raising money at a lower level than $100,000. For example, at Northwestern, our staff of around 30 development officers easily covers the top 500 to 700

TABLE 1. Variables Used in the Model

Variable Name	Description (Dummy Coding 0/1)
CURR 10K	Giving greater than $10,000 in the current 3-year period (1988, 1989, 1990). This is the dependent variable.
PAST 0–99	Average past giving (total pre-1988 gifts divided by number of years since Northwestern graduation) less than $100.
PAST 1+	Average past giving over $1.
PAST 100–499	Average past giving between $100 and $499.
PAST 500+	Average past giving $500 or more.
PAST 3K+	Average past giving $3,000 or more.
RECENT	Donor of a recent gift (1987).
PLANNED GIFT	Planned giving mailer respondent at any time in the past (Dealing-In-Futures or DIF is sent 3–4 times per year and covers topics such as bequests, unitrusts, annuity trusts, etc.).
HIGHPOS	High-position prospect (self-reported CEO, chairman of board, president, etc.).
POTEN HIGH	High gift potential ($250,000+ lifetime to any cause) based on peer review and screening done as part of Northwestern's National Resource Program.
SALARY HIGH	Self-reported salary over $200,000.
SALARY MED	Self-reported salary between $100,000 and $199,999.
PRIZM	Lives in any of the three selected PRIZM clusters: Blue Blood Estates, Urban Gold Coast, or Grey Power.
YEARS 31–40	31 to 40 years since last Northwestern degree.
YEARS 41–50	41 to 50 years since last Northwestern degree.
YEARS 51+	51+ years since last Northwestern degree.
PARTICIPANT	Participant in nongift club organizations.
LOCAL ADDRESS	Resident of local Chicago area.
RELIG	Participated in a student religious activity.

prospects at the $100,000 level. The predominance of past giving to predict these donors also implies that this population is substantially known to the development operation. A model that helps to select the next level down is needed to assist with those individuals who might not be as familiar to the staff, but who are excellent prospects. In the test reported later in this paper, the new $10,000 model with interactions was shown to be superior to the $100,000 model at all levels involving 600 or more potential donors.

Variables

The variables in Table 1 were used in our new model. Some of the coding levels were changed when going from the CHAID results to the logit model to provide consistency among the variables across the three submodels.

Our variables can be thought of as consisting of two types: variables representing individual capacity to make a gift and variables measuring interest in Northwestern. Thus our model crudely represents donor giving behavior as a process in which capacity and interest combine into an actual gift. Variables measuring an individual's capacity to make a large gift include employment position (HIGHPOS), salary, rated gift potential, and the type of neighborhood in which an individual lives (that is, PRIZM cluster). Years since graduation might also capture capacity if we assume that individuals change over the life cycle, with older individuals having greater ability to give since they are much less likely to have the financial responsibilities for children or a new house. Variables best thought of as measures, albeit in some cases indirect measures, of possible interest in Northwestern would include participation in the university clubs, being a resident of greater Chicago, and having belonged to a religious organization while an undergraduate at Northwestern. Finally, past giving, recency of past gift, and interest in planned giving are measures of both capacity to give and interest in Northwestern.

We arrived at the above list of variables by starting with those variables that are theoretically important for predicting donors and then testing a wide variety of models using these variables against our data. Possible interactions were recommended by the results from the CHAID analysis and earlier research results. Categorical instead of continuous variables were used for years since degree and past giving, because of the limited computational capacity of our computer system and the need to allow for interactions. CHAID uses only categorical variables to find interactions. One of CHAID's features allows for the collapsing of categories to those where the largest statistical difference is found.

CHAID and Logit Analysis

We used a combination of CHAID and logit analysis to develop the model of donor behavior. As described in the SAS documentation (SAS Institute, 1991), CHAID partitions categorical data into mutually exclusive subsets that best distinguish between different categories of the dependent variable (using a chi-square test, $p = 0.05$). First, the predictor categories are merged into compound categories such that there are significant differences among the composite categories but not within each one. After this, each independent variable is analyzed and the one that explains most of the variation in the response variable is used to divide the data into subsets based on the values of its merged categories. For each data partition not yet analyzed, the program returns to the merging process. Data partitions that have fewer than a specified number of observations are excluded from further analysis. CHAID is considered an exploratory technique because the selected predictors may be used in further analysis. Once the significant variables are found from among those suggested by theory and

past research, they can be used in a more refined analysis of the data using logit analysis.

Logit analysis is a relatively new technique that is a generalization of regression analysis. It specifies a nonlinear relationship between the probability of an event occurring (e.g., a gift being made) and the variables used to predict that event. By positing a nonlinear relationship, logit analysis solves the problems of using regression analysis to predict rare events such as gifts over $10,000, and of there being substantial numbers of individuals who have not given at all. For further discussion of logit models see Chapter 7 in Hanushek and Jackson (1977).

An obvious question is why we have not just used simple regression analysis or the more sophisticated tobit analysis, which takes into account, in a statistically appropriate way, the fact that there are a large number of never givers. (For an introduction to tobit models see Maddala, 1983.) Both of these techniques use the precise amount that an individual has given in estimating the effects of different variables. In our logit analysis giving is simply coded as being above or below some threshold. We did try both of the above techniques with our earlier models (Lindahl and Winship, 1992), and found logit analyses to have superior predictive power, presumably because we are attempting to predict a rare event—a large gift.

General Model Form

Our model provides a score for each individual represented by the combined weights for each of the individual's attributes. To simplify the presentation of the main interaction with past giving, the full model was subdivided into three parts. Each variable listed in Table 2 involves an interaction with one of the three levels of past giving. Thus, all of the variables in the table are interacted with past giving.

Once the weights for the formula are established with statistical procedures and scores calculated for each prospect in the institution's database, the score is then converted to a probability for each of the prospects [prob $= e^{score}/(e^{score} + 1)$ times 100]. At Northwestern, the converted scores are displayed on the prospect tracking report (prospect "card") together with the scores from the earlier models and assist the development staff in deciding who to visit.

RESULTS

CHAID was unable to handle the full data set (140,000 records) due to insufficient memory on our current IBM 3090 mainframe. Therefore, the results (Figure 1) were obtained by using a randomly selected group of alumni from the full database ($N = 53,276$). Each box in the diagram represents a split of

the data into partitions based on a chi-squared test. For example, cell (15) contains 779 records of those alumni who gave between $100 and $249 in the past and who had a self-reported salary under $200,000 per year. The probability of someone in this group giving a gift over $10,000 in the three-year period of the study is 1.54 percent. This segment is then split into two groups, depending on whether individuals did not (31) or did (32) return a planned giving reply card. The 23 people in the latter group had a 17.39 percent chance of giving a gift over $10,000.

The CHAID program picks past giving as the main predictor variable. Each level of past giving provides an increasing percent chance of someone giving over $10,000 in three years. At the $0 level, the percent is essentially zero. As past giving increases, the probabilities of making a gift over $10,000 are respectively: $1 to $24—0.05 percent; $25 to $99—0.29 percent; $100 to $249—2.33 percent; $250 to $499—5.88 percent; $500 to $749—13.33 percent; $750 and over—33.14 percent. This result confirms our earlier findings (Lindahl and Winship, 1992) that showed past giving is the most important single predictor of future giving.

Second, Figure 1 shows that different variables are important for higher- and lower-level past givers. At the high end, recency of gift is the variable that splits the data. Those who have given in the far past are less likely to give another large gift. At lower levels of the giving spectrum, the capacity measures provide the next split level (gift potential through the peer rating and screening program or self-reported salary).

Finally, there is a further difference among those that gave between $100 and $249 on average in the past. Those with high self-reported salary (16) are subdivided into those who live in certain PRIZM clusters (ignoring the split of past student religion participation because of small sample size, $N = 1$). The alumni who self-report high salary levels but live in nonpredictive PRIZM clusters have a 0.00 percent chance of giving over $10,000 in three years, while those that do live in the PRIZM areas have a 15.56 percent chance. This points us to the possibility that confirmed (via geodemographics) self-reported salaries are more accurate.

Logit Analysis

Using the results of the CHAID analysis as a guide, logit models were developed for each of the three subpopulations based on past giving levels. The submodels are combined to represent a single logit model with interactions. Table 2 describes the coefficients for each variable along with the standard errors. Dashed lines indicate variables that were found to be insignificant and whose effects were then constrained to zero. As suggested by the CHAID analysis, we allow for different effects at different levels of past giving.

Northwestern University Alumni

RANDOM (53,276 out of 140,000)

$10,000+ Prediction Level Over 1988, 89, 90

FIG. 1.

TABLE 2. Independent Variables—Past Giving Interactions

	($500+)		($100–499)		($0-99)	
	Coeff	Std Err	Coeff	Std Err	Coeff	Std Err
INTERCEPT	−2.81	(.29)	−5.80	(.43)	−10.32	(.48)
YEARS 31–40	—	—	—	—	0.78	(.32)
YEARS 41–50	0.69	(.22)	0.75	(.34)	1.38	(.32)
YEARS 51+	0.62	(.25)	0.96	(.36)	1.75	(.35)
RECENT	1.27	(.28)	0.94	(.39)	0.39	(.25)
PAST 3K+	0.90	(.22)	—	—	—	—
PAST 1+	—	—	—	—	1.47	(.50)
HIGHPOS	—	—	—	—	0.60	(.29)
RELIG	—	—	—	—	1.71	(.76)
SALARY HIGH	0.69	(.23)	—	—	1.90	(.40)
SALARY MED	—	—	—	—	1.21	(.33)
PRIZM	—	—	0.60	(.28)	0.40	(.25)
LOCAL ADDRESS	—	—	—	—	0.40	(.24)
PARTICIPANT	—	—	0.63	(.25)	0.56	(.29)
POTEN HIGH	0.53	(.19)	1.25	(.31)	1.82	(.31)
PLANNED GIFT	—	—	—	—	1.68	(.39)
SALARY HIGH * PRIZM	—	—	1.89	(.38)	—	—
POTEN HIGH * SALARY HIGH * PRIZM	—	—	−1.47	(.64)	—	—
YEARS 31–40 * PLANNED GIFT	—	—	2.47	(.60)	—	—
YEARS 51+ * PLANNED GIFT	—	—	1.86	(.57)	—	—

The coefficient columns in Table 2 indicate the effects of the variables used in the model. Since the categories of the variables are coded as zero-one dummy variables, the relative size of their coefficients can be directly compared within and across each of the three subsets.

The evaluation of the absolute size of logit coefficients is somewhat difficult. The dependent variable is the $ln(p/(1 - p))$, or in words, the natural logarithm of the odds of an outcome. Perhaps the easiest interpretation is to exponentiate the coefficient, giving the increase that a one-unit change in the variable produces in the odds of a particular outcome. Thus a coefficient of 1.5 when exponentiated is approximately equal to 4.5, indicating a multiplicative increase in the odds of an event of this amount. For instance, if the odds were originally 2 to 1, with the shift they would be 9 (or 2 × 4.5) to 1.

In the $500+ past giving submodel, there were no additional interaction

variables used. The maximum-likelihood analysis had a chi-square of 50.59 (49 DF) with prob. = 0.4105, showing a good fit for the model. The variable with the greatest effect was RECENT, which follows the results of the CHAID analysis. This implies that alumni who have given a lot in the far past are less likely to give than those who are recent donors. The other variables show that older alumni over 41 years out of school, alumni who are wealthier, and alumni who had given over $3K on average in the past are all more likely to have made a gift of $10K or more.

The second submodel, based on the $100-$499 average past giving subpopulation, had a maximum-likelihood chi-square of 347.82 (789 DF) with prob. = 1.000, which provides an excellent fit indication. The strongest main effect was POTEN HIGH, which follows the CHAID results. The interactions of both YEARS 31-40 and YEARS 51+ with PLANNED GIFT show that alumni who are older and who responded to a planned giving mailer are more likely to give a gift at the $10,000 level than each separately. In other words, younger planned giving mailer respondents are not as motivated to actually set up a planned gift. The interaction of PRIZM with SALARY HIGH shows that a self-reported salary confirmed by geodemographics is better at predicting donors at this level than is either piece of information taken separately. Finally, the negative three-way interaction of SALARY HIGH, POTEN HIGH, and PRIZM adjusts the model for those prospects that happen to have multiple indicators of wealth. Without this part of the model, some people would be scored higher than their actual giving potential.

The third submodel, based on the $0-$99 average past giving, had a maximum-likelihood chi-square of 454.19 (2240 DF) with prob. = 1.000, which again provides indication of an excellent fit. The variables with the largest effects are those that measure capacity—SALARY HIGH and POTEN HIGH. With the intercept at −10.32, a prospect would have to score on many other attributes besides capacity to get to the same level as someone in the other two subpopulations. There were no further interaction effects in this submodel, beyond the principle interaction of past giving with all other variables.

Testing the Model

Our test involves a comparison of the logit model with interactions with the major gifts model ($100,000 level) previously defined (Lindahl and Winship, 1992). Our metric is designed to estimate the selection effectiveness of each model. The metric we use is the sum of giving (including family giving) between 1988 and 1990 by the top potential donors as ranked by our and other models. In other words, the prospects with the highest donation probabilities as predicted from the two estimated models are compared in terms of gift totals

TABLE 3. Comparison of Major Gifts ($100K+) Model with the Logit ($10K+) Model with Interactions

Total Donations 1988, 1989, and 1990 of Selected Alumni

Number Selected	Major Gifts	New Model	Improvement
100	$11,856,717	$10,535,673	$(1,321,044)
500	19,011,737	17,539,999	(1,471,738)
600	19,654,411	20,252,319	597,908
750	20,267,399	21,526,674	1,259,275
1000	20,892,277	22,614,410	1,722,133
1500	21,914,829	23,823,881	1,909,052

for various levels of selection (i.e., the amount of money given by the top 100, 500, 600, etc. prospects selected). Table 3 indicates the dollars raised from the selected group over the following three years. The larger the number of dollars "raised" for the same number of selected alumni, the "better" the model—in other words, the more efficient the model is at identifying the highest potential donors.

Table 3 clearly shows that the logit $10K+ model with interactions is the superior model after the 600th prospect is selected. Although the major gift prospect model does well at the start (under 550 prospects), using the CHAID model accounts for approximately $1 million to $2 million more at the 600 and up prospect levels.

An additional test was done comparing the logit model with interactions with a logit model consisting of the same independent variables—but without interactions—and with the dependent variable at the $10,000+ level. The new model was superior at every test level. Again, the improvement was over $1 million by the 1,500-prospect selected level. Such an increase would certainly justify any additional costs involved in doing a statistical analysis. Because of the high cost of personal visitation, every 100-200 prospects would cost the equivalent of one full-time development officer.

When would it be worth the effort of the development staff to actively pursue a prospect who is in the last subpopulation ($100 > average giving)? In general, the probability of someone in this category giving a gift of $10,000 or more is less than 1 percent. However, looking more closely at those with the highest probabilities shows a different story. Within the top 1,000 prospects with the logit (10K+) model with interactions, there are around 50 prospects from the lowest past giving population. The probability levels for this group of 50 range from 5.8 to 34.8 percent, certainly warranting consideration in the solicitation process. Based on earlier research (Lindahl, 1991), these 50 individuals with low past giving levels are most likely planned giving prospects.

APPLYING THE MODELS TO FIND NEW PROSPECTS

The logit 10K + model with interactions we have developed is about to be utilized at Northwestern for the identification of new prospects. Although we have extensively used only the models developed last year, the latest model has just recently become available to development officers.

The implementation concepts are similar to the other models. Each alumnus on the computer system is scored using the model and the score converted into a probability value. Development officers could then select prospects based on the probability of their giving a gift over $10,000 and choose to visit or call the top prospects. They could look for opportunities to thank the top prospects, beyond standard gift acknowledgment procedures, for whatever size gift is given. When setting up a trip in a certain part of the country, they could use the probability score from the model together with other data to determine who to visit and cultivate for a future gift.

FUTURE RESEARCH

This modelling effort represents a fine-tuning of the earlier models. With past giving playing such a major role in predicting future gifts, the question of how to balance the old prospect pool with new prospects is important. Based on the results from this study, most donors of large gifts will be repeat donors. However, using a logit model with interactions will allow selection of good prospects who are not strong past donors. Our emphasis here has been on predicting which individuals are likely to be donors, not on developing a causal model of donor behavior. Our research, however, does suggest, not too surprisingly, that both an individual's capacity and interest in an institution are critical to their making a gift. It would be desirable to know more about what causes donor behavior, not just what predicts it.

In our analysis past giving is the key predictor of future giving. Is it appropriate to think of past giving as causing future giving, or is it simply a good measure of an individual's latent capacity or philanthropic inclination? Put another way, is past giving simply a good proxy for other variables we have omitted from the model because we simply do not have measures of them or does past giving actually affect future giving? Answering this question would take a different type of analysis than we have carried out. Specifically, an instrumental variable estimation would be needed to correct for the potential correlation between past giving and variables that are omitted from our model. In particular, the effect of past giving may well capture differences across individuals in income and assets, general philanthropic commitment, and the costs of giving not captured by the other variables we have included in our model. Separating out the "true" effect of past giving from its effect due to its role as a proxy for other variables is something we hope to do in the future. We would not be at all surprised to find that past giving had little or no "true" effect on current giving, indicating the lack of any of true structural or behavioral effect. This, however, in no way would reduce the importance of past giving as a predictor variable for indicating which individuals are likely to make gifts in the future.

Gifts, also, are not made in a vacuum. Donors are often visited and solicited by development officers. What are the effects of such visits? If development staff had approached those with the highest capacity (regardless of past giving) over the past few years, would the results have been better than if the staff had approached alumni based primarily on past giving levels? These are questions we are examining now. So far, our results to date (Lindahl, 1991) suggest that the type of gift (outright versus planned) is critical. A measure of development effort will be incorporated into our next models.

REFERENCES

Brittingham, B. E., and T. R. Pezzullo (1990). *The Campus Green: Fund Raising in Higher Education*, ASHE-ERIC Higher Education Report No. 1. Washington, DC: School of Education and Human Development, The George Washington University.

Connolly, M. S. and R. Blanchette (1986). "Understanding and Predicting Alumni Giving Behavior." In J. A. Dunn, Jr. (ed.), *Enhancing the Management of Fund Raising. New Directions for Institutional Research*, no. 51. San Francisco: Jossey-Bass.

Hanushek, E. A., and J. E. Jackson (1977). *Statistical Methods for Social Scientists*. New York: Academic Press.

Hunter, T. W. (1975). What motivates large gifts? Washington DC: Brookings Institution.

Kass, G. V. (1980). An exploratory technique for investigating large quantities of categorical data. *Applied Statistics* 29: 119–127.

Lindahl, W. E. (1991). Differentiating planned and major gift prospects. *Connections*, Westmont, IL, American Prospect Research Association, Fall Issue.

Lindahl, W. E., and C. Winship (1992). Predictive models for annual fundraising and major gift fundraising. In *Nonprofit Management and Leadership*, vol. 3, no. 1. San Francisco: Jossey-Bass.

Maddala, G. S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

Melchiori, G. S. (ed.) (1988). *Alumni Research: Methods and Applications*, no. 60. San Francisco: Jossey-Bass.

Panas, J. (1984). *Mega Gifts: Who Gives Them, Who Gets Them*. Chicago: Pluribus Press, Inc., pp. 141–151.

SAS Institute, Inc. (1991). *SAS Software Application: Using CHAID for Exploratory Analysis of Categorical Data*.

Vickrey, W. (1975). Private philanthropy and public finance. In *Altruism, Morality, and Economic Theory*, ed. Edmund Phelps. New York: Russell Sage Foundation, p. 157.

Weisbrod, B. (1988). *The Nonprofit Economy*. Cambridge, MA: Harvard University Press.