[[1]](#footnote-1)Reunions and Donations. Is there a causal relationship between the two?

By Lasha Bezhashvili,

May 14th, 2018

*Alumni donations constitute an important source of funding for many American universities and colleges, which enable them to offer better equipment, infrastructure, educational programs and other amenities for their students. Given the importance of donations, it is not surprising that many educational institutions try to keep the connection between them and the alumni tangible. One of the ways a college can maintain such connections between itself and alumni and between the alumni themselves is to organize reunions and Davidson College is not an exception. This paper tried to answer the question whether attending a reunion causes the attending alumnus to decide to donate and how demographic patterns influence such a decision. I found that there is a significant causal relationship between the attending a reunion and deciding to donate and that demographic characteristics, such as race, gender, income and involvement on campus are important variables that determine the subsequent donation behavior.*

Davidson College can say for sure that it has produced very generous alumni. According to the data provided by the Office of Giving, the total value of donations that were given to the institution throughout the years from 2012 till 2017 is equal to almost $141,918,000 and the average value of a donation made throughout these years equaled to $918. The trend throughout the past years has also been a positive one: from 2012 to 2017, the total value of donations increased from $17,466,719.04 to $29,432,007.77. This means that there was a 68% increase in donations made throughout such a small amount of time. The reasons behind why an alumnus may decide to donate are various, ranging from family legacies and traditional affinity towards the college to a desire to have one’s name recognized. Finding different reasons of causation could help answer important questions such as: how different are alumni behavioral patterns across different demographic groups? Was it mainly the positive undergrad experience that caused the alumni to donate or other reasons? Are Davidson programs designed to reach out to alumni financially justified? In this paper I tried to find out how much of the decision to donate or not was caused by taking part in one of such programs that are facilitated to keep the connections between the alumni and the college alive: reunions.

**Literature Review**

The literature regarding alumni donations is rich and has provided me with great help on how to construct optimal regression models, which variables to include and how to modify them to fit the regression models better. The first paper that I would like to mention would be “Alumni Donations and Colleges’ Development Expenditures: Does Spending Matter?” by authors: William B. Harrison, Shannon K. Mitchell and Steven P. Peterson. The work was published by American Journal of Economics and Sociology in 1995.

**Data and Descriptive Statistics**

Data was provided by the Office of Annual Giving at Davidson College with the help of its director, Caitlin James. It includes all alumni who were or still are on the mailing list starting from the year of 1840 and contain information about an alumnus’ class year, the most current ZIP code, race, gender, involvement in Greek Life organizations and varsity sports, country of origin and the amount donated during each year starting from 2012 to 2017. I have modified the data to fit it more optimally into my regression models and will explain the names of the independent variables in my regression models and the way I modified the initial data in the following paragraphs.

I will begin by clarifying the dummy variables which indicate the race of an alumnus. I have the following 7 race/ethnicity categories: “white”, “black”, “asian”, “hispanic”, “other”, “missing race” (which is omitted in all my regression models to avoid the dummy variable trap) and “international”. The initial data contained separate categories for biracial, Alaskan/Native American and Pacific Islander races and ethnicities and was inconsistent when it came to international students, because some international students were under the category of “International Student”, while others were under the categories of “White/Caucasian/Non-Hispanic”, “African American”, “Asian American” and others. I decided to make every international student fall under the category of “international” to make the data more consistent. When it came to the categories that were not white, black, asian, hispanic or international, I generated a new dummy variable “other” and made these categories fall under it. The reason behind this decision was the meagre amount of people who were under these categories. Besides the “other” dummy variable, I generated another dummy called “missing race” which contains every US alumni who have not indicated their race or ethnicity. The number of alumni who have not indicated their race or ethnicity was substantial and dropping them would mean losing a significant part of my data. No alumnus can fall under more than 1 category.

The alumni in the data set fell under 2 gender categories: male and female. I generated a dummy variable called “female”, which is equal to 1 if a given aluminum is indicated as female and 0 if it is indicated as male. Similar logic applies to the independent variables of “reunion2014”, which is equal to 1 if an aluminum attended the reunion in 2014 and 0 if not, “pcourt”, which is equal to 1 if an aluminum was part of a fraternity/sorority or an eating house and 0 if not, “sports”, which is equal to 1 if an aluminum participated in varsity sports and 0 if not.

I modified the data regarding ZIP codes. The number of ZIP codes themselves do not explain anything, but can function as proxies for income, and given the fact that the paper is concerned with donations, having information about incomes levels is essential. I have matched the alumni ZIP codes with the median household incomes for the relevant ZIP codes. The data for median household incomes was downloaded from the website of The University of Michigan's Population Studies Center. The independent variable “medincome” indicates the median household income of the ZIP code for a given alumnus. I also divided the alumni into 4 quartiles based on the amount of their median household incomes of their ZIP codes. The “quartile1” dummy variable is equal to 1 if an alumnus falls under the first quartile and 0 if not, “quartile2” is equal to 1 if an alumnus is between the 1st and 2nd quartiles and 0 if not, “quartile3” is equal to 1 if an alumnus is between the 2nd and 3d and 0 if not and finally, “quartile4” is equal to 1 if the alumnus is above the 3d quartile.

I also modified the data about the graduation year. I made a new variable called “yearsfter” which measures how many years ago an alumnus graduated as of 2018.

The dependent variable of all my regression models is “donation”, which is equal to 1 if an

alumnus has donated any amount throughout the years of 2014 till 2017 and is equal to 0 if not.

**Econometric Models and Results**

Donations from alumni make up a large portion of funding received by American colleges and universities and Davidson College is not an exception. Given its importance, many programs are facilitated to keep the connections between the alumni and the college alive. The participation in these different programs and memberships in different giving societies give the alumni recognition of their name and of their contributions to the development and advancement of the educational institution, while colleges and universities receive increased revenue as a reward for recognizing the alumni (Harrison et. al). It should also be said that name recognition is not the only reason to why alumni donate, for some of them it can be a way to express gratitude for the education received, while for others, usually in the case of legacies, donation may be a family tradition. My initial objective in this paper was to find whether Davidson’s membership programs increased the chances of alumni donating and what were their causal effects on the probability of them donating, but given the limitations of the data I had, I was unable to construct regression models that would enable me to answer the question. The reason to why the data was insufficient was because I did not have information about whether the alumni donated or not during the years before these giving societies were introduced. For example, the Ne Ultra was launched in 1972, while I only have data about donations throughout the years of 2012 till 2017. Given this limitation, I am unable to find out whether the introduction of the Ne Ultra Society had a statistically significant effect on alumni donation patterns at Davidson College.

Since I had to change my initial objective, I decided to switch my attention to reunions. I had the data about whether the alumni in my data set have attended the reunion in the year 2014 or not. Given the availability of this information, one could regress the outcome of whether the author donated during the years of the years after 2014 on them attending the reunion in 2014, but given the fact that reunions are organized every year by Davidson College and I only have information whether the alumni have attended the one in 2014, there will be an omitted variable bias which will probably overestimate the coefficient on the variable of interest, since an alumni may have decided to donate after attending the reunion in 2015 rather than the one in 2014. The other reason to inconsistency of my coefficients could potentially stem from errors in variables. I don’t have information whether the aluminum donated before the reunion happened in 2014 or after, so potentially there could be an alumnus who donated before the reunion was organized and decided to attend, but in my regression model I implicitly assume that attending the reunion had a causal effect on whether he or she donated or not.

As described in the paragraph above, there are very important threats to validity, but even though I am unable to retrieve the data on reunions that happened during other years and therefore find immediate remedies to the threats described above, I still decided to go on with this question and try my best to come close to answering the question whether there is a statistically significant effect of alumnus attending reunions on whether an alumnus decides to donate or not. The answer to this question could indicate whether organizing and bearing costs of organizing unions are financially justified (this does not mean that I think that the reason behind organizing reunions should be trying to get more donations out of the alumni). Answering the question whether there is a causal relationship between my variable interest and donations could begin by a naive linear probability regression model that would have the following form:

The coefficient I for the variable of interest is statistically significant at the 1% significance level and is equal to 0.390. This means that an aluminum that attended the reunion in 2014 is 39% more likely to donate than an aluminum who did not. The analysis cannot obviously stop at this point, the probability of whether the alumnus donates or not is determined by many other variables which are also potentially correlated with whether the alumnus attends the reunion or not. The absence of these independent variables in our model causes omitted variable bias which makes our coefficient on our variable of interest inconsistent. An example of an independent variable which causes omitted variable bias in our regression model above could be whether the alumnus was involved in Greek Life or not. According to the paper “Alumni Donations and Colleges’ Development Expenditures: Does Spending Matter?” by William B. Harrison, Shannon K. Mitchell and Steven P. Peterson, being involved in a fraternity or a sorority increases the amount of money donated by 2% and the coefficient was statistically significant at the 1% significance level. The data he utilized to construct his regression models were gathered from eighteen different colleges and universities in the US and included both large and small, public or private universities. Since he utilized data from small private colleges, there are grounds to assume that his results are externally valid and can be applied to Davidson too. One can also assume that there should be correlation whether an alumnus attended the reunion or not with whether he or she was part of the college Greek Life, the main reason behind this assumption being the tendency of members of such organizations to be more connected to the college after graduation compared to those who were not members. Since we determined that being in a fraternity has a statistically significant effect on donations and that it is correlated with whether an alumnus attends a reunion or not, not including the variable should cause omitted variable bias. In my following model, I added the variable “pcourt”, which is equal to 1 if the alumnus was part of a Patterson Court (Greek life) organization and 0 if not:

As expected, the coefficient on our variable of interest decreased since it captured the effect of being in a fraternity, sorority or an eating house on the dependent variable. The coefficient on the variable “pcourt” is statistically significant at 1% level. According to the second regression model, an alumnus who was involved in Patterson Court organization is 20% more likely to donate than an alumnus who was not, keeping all else equal. The R squared of the second model is equal to 0.0657, while the R squared of the first model is 0.0241. This means that the second model fits the data provided better than the first one and therefore is superior with predicting our dependent variable.

    Similar to the reasoning behind introducing the “pcourt” independent variable, I decided to introduce variables which control for the median household income for the ZIP code the alumni currently reside in, time passed since graduation, gender, race, income and athletics and utilize different regression estimation models and forms. The results are summarized in the table below:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | |  | (1) | (2) | (3) | (4) | (5) | (6) | (7) | | VARIABLES | donate | donate | donate | donate | donate | donate | donate | |  |  |  |  |  |  |  |  | | reunion2014 | 0.391\*\*\* | 0.364\*\*\* | 1.238\*\*\* | 1.239\*\*\* | 1.065\*\*\* | 1.056\*\*\* | 1.057\*\*\* | |  | (0.00923) | (0.00965) | (0.0933) | (0.0936) | (0.0888) | (0.0894) | (0.0895) | | yearsfter |  |  | -0.0322\*\*\* | -0.0320\*\*\* | -0.0134\*\*\* | -0.0125\*\*\* | -0.0124\*\*\* | |  |  |  | (0.000623) | (0.000625) | (0.00104) | (0.00107) | (0.00108) | | quartile2 |  |  |  | 0.175\*\*\* | 0.131\*\*\* | 0.130\*\*\* | 0.129\*\*\* | |  |  |  |  | (0.0342) | (0.0392) | (0.0402) | (0.0402) | | quartile3 |  |  |  | 0.238\*\*\* | 0.207\*\*\* | 0.202\*\*\* | 0.202\*\*\* | |  |  |  |  | (0.0348) | (0.0387) | (0.0395) | (0.0395) | | quartile4 |  |  |  | 0.233\*\*\* | 0.224\*\*\* | 0.215\*\*\* | 0.215\*\*\* | |  |  |  |  | (0.0348) | (0.0399) | (0.0407) | (0.0407) | | pcourt |  | 0.204\*\*\* | 0.250\*\*\* | 0.252\*\*\* | 0.269\*\*\* | 0.260\*\*\* | 0.234\*\*\* | |  |  | (0.00602) | (0.0252) | (0.0252) | (0.0277) | (0.0284) | (0.0334) | | sports |  |  | 0.0816\*\*\* | 0.0823\*\*\* | 0.0479 | 0.0436 | 0.0443 | |  |  |  | (0.0268) | (0.0268) | (0.0295) | (0.0301) | (0.0301) | | white |  |  | -0.170\*\*\* | -0.168\*\*\* | 0.0979\*\* | 0.115\*\*\* | 0.116\*\*\* | |  |  |  | (0.0324) | (0.0324) | (0.0400) | (0.0405) | (0.0405) | | black |  |  | -0.740\*\*\* | -0.744\*\*\* | -0.375\*\*\* | -0.357\*\*\* | -0.347\*\*\* | |  |  |  | (0.0802) | (0.0804) | (0.0799) | (0.0802) | (0.0804) | | asian |  |  | -0.608\*\*\* | -0.599\*\*\* | -0.207\* | -0.184\* | -0.183\* | |  |  |  | (0.108) | (0.108) | (0.108) | (0.108) | (0.109) | | international |  |  | -0.906\*\*\* | -0.853\*\*\* | -0.568\*\*\* |  |  | |  |  |  | (0.0688) | (0.0700) | (0.0717) |  |  | | hispanic |  |  | -0.613\*\*\* | -0.605\*\*\* | -0.185 | -0.162 | -0.160 | |  |  |  | (0.118) | (0.118) | (0.120) | (0.120) | (0.120) | | other |  |  | -0.298\*\* | -0.297\*\* | 0.188 | 0.215 | 0.217\* | |  |  |  | (0.130) | (0.130) | (0.131) | (0.131) | (0.131) | | female |  |  | -0.00412 | -0.00208 | 0.0809\*\*\* | 0.101\*\*\* | 0.0502 | |  |  |  | (0.0315) | (0.0315) | (0.0309) | (0.0319) | (0.0472) | | medincome |  |  | 2.73e06\*\*\* |  |  |  |  | |  |  |  | (4.25e-07) |  |  |  |  | | interaction |  |  |  |  |  |  | 0.0886 | |  |  |  |  |  |  |  | (0.0608) | | Constant | 0.523\*\*\* | 0.425\*\*\* | 1.280\*\*\* | 1.284\*\*\* | 0.629\*\*\* | 0.595\*\*\* | 0.602\*\*\* | |  | (0.00318) | (0.00429) | (0.0542) | (0.0519) | (0.0626) | (0.0640) | (0.0642) | |  |  |  |  |  |  |  |  | | Observations | 25,772 | 25,772 | 15,088 | 15,088 | 11,281 | 10,819 | 10,819 | | R-squared | 0.024 | 0.066 |  |  |  |  |  | | PseudoR2 |  |  | 0.305 | 0.306 | 0.0824 | 0.0723 | 0.0724 | | Robust standard errors in parentheses | | | | | | | | | \*\*\* p<0.01, \*\* p<0.05, \* p<0.1 | | | | | | | | |

The third model utilized all the independent variables in my data set and it’s a probit regression model. I decided not to use the linear probability model after the second regression model because of two reasons. The first reason is because predicted variables calculated through LPM can exceed the value of 1 and therefore are nonsensical. Given the fact that I introduced more independent variables, which increased the probability of having more extreme values that can yield more dependent variables that are nonsensical, using a probit model would provide more sensical predictions of how likely the alumni are to donate or not, even with extreme values. The second reason to why I decided to use the probit model is because it captures the non-linear relationship between the age and income of an alumnus with the probability of whether he or she donates or not. The theory behind to why I believe there is a non-linear relationship between age and our dependent variable comes from the “Alumni Donations and Colleges’ Development Expenditures: Does Spending Matter” paper by authors mentioned in the previous paragraph. According to the source, the growth rates of alumni donations tend to increase as the alumnus ages but begins to decline after the age of 52. Because of external validity of their regression models, I decided that this pattern would hold true for Davidson College as well. Regarding income, I concluded that up to a certain point, a person does not have enough disposable income to make donations, but after that point is passed, the rate of increase of probability whether the person donates and the size of his or her donation should begin growing at an increasing rate.

The fourth regression model uses the same probit model, data and independent variables except for “medincome”. The coefficient of “medincome” is statistically significant at the 1% level but its coefficient is very small and in a way, difficult to interpret. I decided that instead of using the contiguous variable, I would use the quartile dummy variables which divide the alumni into 4 income categories and run the regression again. The coefficient on other variables either did not change or barely changed at all and the values of the pseudo-R squared are almost identical in both models. The coefficients on probit models cannot be numerically estimated directly, so it would be necessary to provide marginal effects of the change in the variable of interest or of change in other variables to paint a clearer picture of the outcomes of the regression models. For example, if we take a male white American alumnus whose projected median household income is in the 2nd quartile, who was involved in varsity sports while at Davidson College and who graduated 5 years ago and compare it to the alumnus with the exact same characteristics besides one attending the reunion and the other not, the one who has attended the reunion is 10% more likely to donate. If we make the exact same comparison but change the race of the 2 alumni being compared from white to black, the difference increases to 23%. I also decided to look how being involved a Patterson Court organization impacts whether an alumnus donates or not. I decided to compare a male white American alumnus who graduated 8 years ago, whose projected median household income is in the 2nd quartile and who did not attend the reunion in 2014 and compare to an alumnus with the same exact characteristics except for the involvement in the college’s Greek Life. Turns out, the alumnus who was involved in the Greek Life is 4.6% more likely to donate. If we make both alumni attend the reunion, the difference shrinks to 0.5%, if we make the one who was involved in Greek Life attend the reunion and the other not, the difference increases to 12%.

In my fifth regression model, I kept the model the same as the previous one, but modified my data. I dropped every alumnus who graduated more than 60 years ago to ease the bias introduced through errors-in-variables. The alumni who graduated more than 60 years ago tend to have the information about their race missing more compared to more recent Davidson graduates (probably because Davidson College back then used to be almost exclusively white and therefore saw no need in collecting data about race) and their estimated median household incomes by ZIP code are probably very far from reality, because the alumni dropped are probably retired and don’t have income any more. I decided to choose an arbitrary cut-off rather than selectively dropping alumni without information about their racial or ethnic background in order to avoid introducing any correlation that could arise from the “filtered” alumni and the error term and to avoid having a biased sample (for example and on purely theoretical level, the people who do not wish to have their race specified or reported may share ideas that are associated with liberalism, which itself is correlated with higher income and dropping them would therefore bias my data towards lower-income alumni). The most substantial change in coefficients compared to those of the previous model can be observed in the “female” and “white” variables. The independent variable which accounts for whether the alumnus is a female or not has become statistically significant at the 1% significance level and the relationship between “white” and the dependent variable changed from negative to positive. The reason behind the change in the “white” independent variable is easier to establish. As noted before, many of the alumni who graduated 60 years and earlier tend to have their information missing and were grouped under the “missing race” category, while being white in most of cases. As a result, the “missing race” variable captured the effect of being white. The result with the female is a more interesting one, since it runs contrary to some of the literature. According to the paper “Who Are the Alumni Donors?” by Charles T. Clotfelter “men and women were about as likely to make any gift” (Clotfelter 129), but in our case, being a woman increased the probability of donating and its coefficient is statistically significant. If we look at the average marginal effects

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |

1. I would like to thank Caitlin James of the Davidson College Office of Giving for providing and clarifying the data that was necessary to conduct this project. [↑](#footnote-ref-1)