INTERTEMPORAL AND CROSS-SECTIONAL SUBSTITUTION IN CHARITABLE GIVING

An Undergraduate Research Scholars Thesis

by

DEREK WELBORN

Submitted to Honors and Undergraduate Research Texas A&M University in partial fulfillment of the requirements for the designation as

UNDERGRADUATE RESEARCH SCHOLAR

Approved by Research Advisor: Dr. Jonathan Meer

May 2013

Major: Economics

TABLE OF CONTENTS

Page

TABLE OF CONTENTS……………………………………………………………………..1

ABSTRACT..2

CHAPTER

I INTRODUCTION..3

II METHODS..6

III RESULTS... 10

IV CONCLUSION 17

REFERENCES 18

TABLES 19

ABSTRACT

Intertemporal and Cross-Sectional Substitution in Charitable Giving. (May 2013)

Derek Welborn

Department of Economics

Texas A&M University

Research Advisor: Dr. Jonathan Meer

Department of Economics

This paper focuses on two aspects of matching a project. Using aggregate data obtained from an online fundraising platform, DonorsChoose.org, Using a fixed effects regression model to control for the time invariant and unobserved characteristics of matches, we examine the effect that matching a project has on the project receiving a donation . We also aim to find out the effect that matches have on the amount of money of projects that have received funding. We expect our results to show that matching a project increases the chances of it receiving a donation. We also expect that the match has a negative effect on the amount of money raised because would be matchers modify their donation depending on the type of match. Our results support the theory of substitution between different charities.

CHAPTER 1

INTRODUCTION

This paper aims to examine the effect that matching a donation has on the amount of money raised by a charity across several time periods. Economists have long been interested in what makes people donate their money. In 2009 alone, nearly 304 billion dollars was given to charitable causes.

Most of the existing literature on the topic has been on the effect that matching a donation has on the amount of money raised. (Karlan and List 2007 & Huck and Rasul 2009). The field experiments conducted in the two papers have shown that linear matching systems –that is donating a certain dollar amount for every dollar contributed by donors result in an increase in the overall donations received but a decrease in the amount contributed by the donor as compared to non-matched donations.

The field experiments above have examined only one charitable cause each. The effect that a certain match has on similar causes across different time periods has not been examined due to the nature of the experiments performed in previous studies.

We obtained data from DonorsChoose, an online fundraising platform for public school teachers that acts as an intermediary between teachers across the United States who need funding and the donors. The data consists of aggregate data of the projects (each project corresponds to a certain teacher that needs funding), matches and donations from 2007 to 2012.

We will be able to see the effect that a match has on the whole environment of charitable giving. For example, when a certain donor decides to match a dollar for every dollar contributed to high school math classes in Texas, from previous literature we expect the number of donors and overall donations to increase. But where do the increased numbers of donors come from? Some of them may be first time donors but some of them will also be people who usually donate to other causes but donated to this one due to the presence of the match. When the match expires the following month and a new project for high school math classes in Texas goes up will it be less likely to receive funding due to the previous month's match?

We will be able to answer those questions with the data. We will carry out statistical tests for significance and run two separate regressions. The first with the amount per day received by the charity as our variable of interest which will help answer our first question and the second with the probability of getting funded as our variable of interest which will help answer our second question.

Our first question was how the presence of a match affected the the probability of receiving a donation. The response variable was the amount of money received per day by a project. The variable of interest for our first question was the match itself. The right hand variables in our regression were the number of days since the match began, the number of competitors; all matched projects on a given day and all projects on a given day and the number of days and months since a project was last eligible for a match.

Our second question dealt with the impact of matching and whether matching actually brings in new dollars or just shifts the dollars from other projects (or a later time period) to the matched projects. In order to answer that question we had to examine how money flowed into various projects when matches were live. We could then compare the money flow for the matched projects to similar projects that were unmatched at the same time period as well as past and future time periods.

CHAPTER 2

METHODS

We obtained 3 data sets from DonorsChoose: a projects data set, a donations data set and a matching data set. The data sets contained data covering a 10 year period from 2003 to 2012.

The projects data set consisted of an observation for each project in the 10 year period, the project's characteristics (such as location, subject and grade level), whether the project was matched and the start and end dates for when the project was live (when donors could donate to that project). Originally we started with 438,234 observations. We dropped 36,452 from 2002 to 2006 because the data from those years contained incomplete information regarding the project characteristics. We dropped a further 14 observations because they contained negative project costs and we dropped 1 more observation because it took -1 days to fund. We ended up with 401,677 observations.

The donations data set consisted of an observation for each donation made in the 10 year period, its characteristics (amount, cash/credit card and donor location) and the day on which the donation was made. We dropped 32,459 observations from the data set because we had dropped the projects which the donations were made to in the projects data set. We ended up with 1,770,254 donations.

The matching data set consisted of an observation for each match from 2008 to 2012, its characteristics and a start and end period for which the match was live. We had 1,303 matches, each of which could potentially have had numerous projects which met its characteristics. For example, if corporation A matches high school math classes in Texas for March 2012, any high school math class in Texas, in March 2012 was eligible. However, meeting all the eligibility requirements does not guarantee that a certain project will be matched. Donor's Choose decides which eligible matches will actually get matched. Donor's Choose implements two types of matches:

1. Double Your Impact: A 1:1 linear matching scheme in which the corporate or foundation partner donates one dollar for every dollar donated by the public.

2. Almost Home: A matching scheme in which the corporate or foundation partner donates the last $100 of an almost funded project.

Using Stata, a data analysis and statistical software, we merged the project data set and the donations data set since they had a unique identifier, the project id, in common. Each project had a start date and an end date and each donation had a date on which it was made. After merging the two data sets we ended up with a panel in which each project had an observation for each donation that was made to the project.

We then coded the matches onto the merged project-donation data set so that for each project we had its various characteristics, the donation characteristics, the day the donations were made and all the matches that the project was eligible for. Having the matches that it was eligible for even though it was not matched would provide a control group for the effects of matching. For example comparing the money raised by two math high school classes when both were eligible for a match and only one was matched.

We had all the matches that a project was eligible for by criteria only, not by date. The dates provided to us to Donor's Choose were inaccurate and were just a rough start and end point to the dates.

There were 332 matchers; a matcher was defined to be the donor who contributed to the match. The 332 matchers combined for 409 matches. 55 of the 332 were donors who matched more than one project. If a project met the criteria it was matched. Each of the 409 matches was further divided into match criteria id of which there were 1,303. Match criteria ids were subsets of the matches themselves they met all the criteria of a match but had extra criteria to distinguish themselves from other match criteria id. An example of a match would be donating to all third grade math classes, which met certain poverty thresholds in Texas. A third grade math class in Austin that met those criteria may have a different match criteria id than a third grade math class in Dallas even though both are part of the same match.

Using the donations data we individually isolated all the donations made by each of the 332 donors from the donations data. For the 277 donors who had one match we took the first match they made to be the start date. For the 55 multiple time matchers, we sorted the donation dates for each individual, deleted the donations made on the same day and created a variable that counted the days between donations. In most cases the days between donations would range from 0-12 then there would be an observation with a days since value in the hundreds we assumed that to be the start date for other matches the donor started. After obtaining all our start dates for every match from each donor, we merged those start dates into the donations file. We assumed that the donation made on the date before the long gap that we assumed to be our start date was our end dates.

We then had a file containing all 400,000 projects, all the matches each project was eligible for (by criteria) and the dates that those matches were live. The file contained 5,419,455 observations.

We then took a random 10 % sample of the number of projects. All the results we infer will come from this 10% sample. Using the 39, 723 projects in our 10% sample we created an observation for each for each day that a project was live and got 3,244,972 observations. Our reason for taking only a 10% random sample was due to tractability. With only a 10% sample we had 3 million observations.

Using the sample we created a number of variables. And for our final sample our main independent variables were as follows.

a. Number of days the project has been live and its square
b. An indicator for whether the project was eligible for a match on that day
c. The total number of other matched projects active on that day
d. The total number of other projects active
e. Days and months since the last eligible match

CHAPTER 3

RESULTS

The effect of a project being matched has an effect on how much money is raised by a project. Matching and lead donor treatments (Huck and Rasul [2007]) have been found to have different effects on the probability of giving and the amount of giving. Therefore, our first variable of interest is the probability of receiving a donation on a given day. Our model took the form

$$(1)\ A_{it} = \beta_0 D_i + \beta_1 D_i^{\ 2} + \beta_2 M_{it} + \beta_3 MatchedProjects + \beta_4 AllProjects + \beta_5 DS_i + \beta_6 DS_i^{\ 2} + \beta_7 TimeEffects + \beta 8 Projecti + u_i$$

Where M_{it} is an indicator equal to 1 if project i received a donation on day t and 0 otherwise. t being the number of days that a project was live. A_{it} represents the amount of money raised by project i on day t. Taking the log of A_{it} allows us to only consider projects that received donations. D_i signifies the number of days that a project i has been live. DS_i signifies the number of days since the last eligible match, with those that never had an eligible match dummied out. M_{it} is a dummy variable that is 1 when the project i was matched on day t. MatchedProjects is a categorical variable that divides the projects that were matched into quintiles based on the number of matched projects that were live were on day t. AllProjects is a categorical variable that divides the all projects on day t into quintiles based on the number of projects that were live on day t. TimeEffects is a categorical variable that gives the effect of each month by year. The $Projects_i$ variable refers to the fixed effect for each project that control for unobserved heterogeneity by controlling for any time invariant effects for each project. It holds the projects unobserved characteristics such as the pictures of the children or the message by the teacher fixed. $U_{i\ refers}$ to the error term.

Our aim is to find out how the probability of receiving a donation by a project is affected by our variables of interest. We begin by examining the effect that the number of days that a project is live has on the probability of receiving a donation by a project. We expect to observe that the longer a project is the more likely that it will get funded for a variety of reasons. The longer a match lasts the more likely it is "seen" by a larger number of people some of whom would make a donation also people who make a donation might encourage their friends in the future to donate to the project.

The number of days effect is positive and significant. Projects that have been live for a longer period of time have a larger probability of receiving a donation on any given day. The length of a project is associated with .02 percentage point more likelihood of receiving a donation. The number of days squared effect is positive and significant. Projects that have been live for a longer period of time have a larger probability of receiving a donation on any given day. The length of a project squared is associated with .0004 percentage point less likelihood of receiving a donation.

Although the days squared coefficient is significant it's negligible. The days squared variable was included to give us insight into how the number of days that a project was live affect the probability of the project receiving a donation. The positive function tells us that increasing the number of days on a project increases the likelihood of getting funded. The negative coefficient on the days squared variable tells us that the effect of increasing the number of days is an increasing function that decreases as time increases.

The days and the days squared variable could be collinear and as such their coefficients would not be significant. We ran a joint f-test for the significance of the 2 variable. The results show that the number of days since a project was posted is highly significant.

We know that the longer a project lasts the more likely it will receive a donation but we are interested in estimating the effect that an extra day has on the probability i.e. the marginal effect at the mean. .

Our results showed that, evaluated at the mean, each additional day the project is live makes it .02 percentage points more likely to receive a donation on a given day and our result is highly significant.

The next coefficient on table one is the match. The match was a dummy variable that was 1 if the project was matched and 0 otherwise The coefficient tells us that projects are .4 percentage points more likely to receive a donation on days they are matched than on days they are not, all else equal. A project has a 2.8 percentage point change of being matched. The coefficient of match is about 15% of our baseline measure. The effect of being matched is significant.

In the original data set we had for each project the total number of matched projects available on each day. We then created a categorical variable that divided the total number of matched projects available into quintiles. The effects of which would give us the effect that the amount of matched competitors for a donation has on the probability of receiving a donation on a given day. The results in table one show that a project with the most matched competitors (18,235 to

24,260) is .8 percentage points less likely to get a gift relative to a project facing the least competition from other matches (0 to 87), all else equal.

Because the quintiles of other matched projects are related and could be collinear we used an F-test to test for joint significance. The results show that the categories are significant.

We then decided to test whether the categories are statistically different from each other in order to see whether our coefficients in table 1 would hold. The results show that the second and fifth quintiles are not statistically different from each other. All other quintiles are statistically different from each other.

From our original data set we had a variable for each project total number of projects (regardless of whether they were matched) on that day. We then created a categorical variable that divided the total number of projects available into quintiles. The effects of which would give us the effect that the amount that all competitors for a donation has on the probability of receiving a donation on a given day. The results in table one show that a project with the most matched competitors (23,131 to 26,982) is .26 percentage points less likely to get a gift relative to a project facing the least competition from other projects (20 to 12,762), all else equal. Therefore the more competition a project receives the less likely it is to receive a donation.

As with the previous set of possibly collinear indicators we ran an F-test to test for joint significance. The results show that the indicators are statistically significant.

The next coefficient in table 1 is daysbef and its square. The daysbef variable measures the days since the last eligible match for each project with projects that never had eligible matches dummied out. As with the number of days that the match was live, the daysbef variable has a positive effect while its square had a negative effect. This tells us that the longer it's been since you were eligible for a match, the more likely you are to get a gift (though at a decreasing rate).

We then tested for joint significance of the daysbef and its square as well as testing for the marginal effect of the each additional day that a project was not eligible for a match as we did with the variable a number of days a project has been live. The results from the two tables show that the daysbef and its square are significant and that, evaluated at the mean, each additional day that has elapsed since the last match for which that project would have been eligible increases the probability of getting a donation by .0035 percentage points. While this is statistically significant, it's extremely small.

We include the number of months since the last eligible for each project with projects that never had eligible matches dummied out. The results show that each additional month that has elapsed since the last match for which that project would have been eligible increases the probability of getting a donation.

Our second question dealt with how much money a project raised on a given day conditional on it getting funded on that day. We modeled our equation as follows.

$$(2)\ Log(A_{it}) = \beta_0 D_i + \beta_1 D_i^{\,2} + \beta_2 M_{it} + \beta_3 MatchedProjects + \beta_4 AllProjects + \beta_5 DS_i + \beta_6 DS_i^{\,2} + \beta_7 TimeEffects + \beta_8 Project_i + u_i$$

$Log(A_{it})$ refers to taking the natural log of the amount project i receives on day t. $Log(A_{it})$ was used because it running a regression with with it as our response variable only runs regressions on projects that received greater than 0 funding all projects that received funding.

The results of our second regression are in table 1.

The number of days effect is positive and significant. Projects that get funded and have been live for a longer period of time tend to receive a higher donation on average. The length of a funded project is associated with 1 percentage increase in the log of the donation amount which corresponds to a 1.107 dollar increase per donation. The number of days squared effect is negative and not significant.

The days and the days squared variable could be collinear and as such their coefficients would not be significant. We ran a joint f-test for the significance of the 2 variable. The results show that the number of days since a project was posted is highly significant.

We know that the longer a project lasts the more likely it will receive a donation but we are interested in estimating the effect that an extra day has on the amount received by a project.

Our results showed that, evaluated at the mean, each additional day the project is live makes increases the log of the donations .98 percentage points on a given day and our result is highly significant which means that each additional day a match is live adds an additional dollar to the amount raised.

The next coefficient on table one is the match. The coefficient was not significant. The five coefficients on the indicator variables for all eligible matched projects that received funding were insignificant as well.

We ran an F-test to test for joint significance. And saw that the variables were not significant. The quintiles of the projects that received donations covered lower ranges and as a result we could not infer from them.

The next coefficient in table 1 number of days since the last eligible match for each project with projects that never had eligible matches dummied out. As with the number of days that the match was live, the number of days since the last eligible match for each project variable has a positive effect while its square had a negative effect. This tells us that the longer it's been since you were eligible for a match, the more likely you are to receive a larger donation.

CHAPTER 4

CONCLUSION

Using a combination of different data sets from DonorsChoose we were able to obtain a unique data set which contained all the projects, the matches they were eligible for and the amount of money they raised for each day. We took a 10% random sample of the data for tractability purposes and the results and conclusions from our paper are based on the 10% random sample.

We have estimated the effect that matching a donation has on the probability of charitable contributions to a project. We find that the effect is very significant which confirms the previous studies and field experiments done regarding the effects of matching on fundraising.

Other factors that were found to affect the probability of receiving a match are, the amount of days that a project has been live, the longer a project is active the more likely it is to receive a donation. Also, competition was found to have a negative effect on receiving a donation as was expected with many projects competing for a limited amount of dollars.

Our results show that it is a worthwhile endeavor for charitable causes to seek matchers because the matches increase the likelihood of receiving a donation. A further question that remains unanswered is whether a match truly brings in more money or just shifts money from projects that would have been funded. We hope to be able to answer those questions with our data set.

REFERENCES

Huck, Steffen and Imram Rasul. "Comparing Charitable Fundraising Schemes: Evidence from a Natural Field Experiment." B.E, Journal of Economic Analysis and Policy, 10(1), Article 31

Karlan, Dean and List, John. "Does Price Matter in Charitable Giving? Evidence from a Large Scale Natural Field Experiment. The American Economic Review, Vol 97, No. 5 (Dec. 2007) pp. 1744- 1793

Reinstein, David. "Does One Charitable Contribution come at the expense of another?" The BE Journal of Economic Analysis & Policy, 2010

TABLES

Table 1: Fixed Effects Regression Results

	Probability of a Project getting funded OLS Regression with Fixed Effects	Conditional on getting funded amount of money raised OLS Regression with Fixed Effects
Number of days that project has been live	0.000***	0.010***
	(-0.0002699)	(-0.0101078)
Number of days that project has been live squared	-0.000***	0
	(-0.000000435)	(-0.00000194)
Whether a project was matched or not on a given day	0.004***	-0.028
	(-0.0043287)	(-0.027924)
First quintile of matched projects	0	0
	(.)	(.)
Second quintile of matched projects	-0.005***	0.023
	(-0.0045574)	(-0.0231818)
Third quintile of matched projects	-0.003***	0.038
	(-0.0027018)	(-0.0383846)
Fourth quintile of matched projects	-0.003*	0.014
	(-0.0034234)	(-0.0135915)
Fifth quintile of matched projects	-0.008***	0.084
	(-0.008243)	(-0.0843184)
First quintile of all projects	0	0
	(.)	(.)
Second quintile of all projects	-0.001	-0.064
	(-0.0010436)	(-0.0637067)
Third quintile of all projects	-0.005***	-0.063
	(-0.004588)	(-0.0628523)
Fourth quintile of all projects	-0.004***	-0.004
	(-0.0041252)	(-0.0039594)
Fifth quintile of all projects	-0.003	-0.044
	(-0.0026606)	(-0.0440135)
Days since last eligible match	0.000***	0.004***
	(-0.0000629)	(-0.0042465)

Days since last eligible match squared	-0.000***	-0.000*
	(-0.0000000931)	(-0.00000317)
_cons	0.045**	7.289***
	-0.0453457	-7.288646