Intuitive Donating: Testing One-Line Solicitations for $1 Donations in a Large Online Experiment[1]

February 2018

Samantha Horn and Dean Karlan

Abstract

We partnered with a large online auction website to test differing messages' effects on the decision to donate to charity at checkout. Our setting, where impulsive decisions are likely to be driving donations, allows us to evaluate intuitive responses to messages prompting a donation. We find that shorter messages, matching grants, and descriptions of a charity's mission increase both the likelihood that a user donates, as well as the average amount donated. Conversely, displaying the impact of the donated amount, the popularity of the charity, and that a charity uses scientific evidence do not improve donation rates. These results contribute to our understanding of how framing requests drives the decision to donate and are practically relevant to the many retail sites which promote giving at point of sale.

JEL Codes: C93, D64, H4
Key Words: Charitable Giving

Introduction

In recent years, many experimental studies have been conducted to study charitable giving decisions, including whether to give, how much to give, and where to give. Typically these studies are done either in a laboratory experiment in which individuals are given information about charities and asked to make a decision, or in a non-laboratory setting, such as direct marketing mail, email, or door-to-door.

Naturally, the channel and decision-making environment may influence charitable giving decisions, not merely in the level of giving but in the rank order of effectiveness of different messages (this is a specific example of the general point made by Deaton 2009). For example, some donation decisions are undoubtedly intuitive, instantaneous, and impulsive (akin to System I decisions in Kahneman 2003), whereas others are deliberative, protracted, and thoughtful (akin to System II decisions in Kahneman 2003). A channel that likely leads to deliberation, such as a laboratory setting in which participants know they are being studied, may yield different sets of results than one that allows only for minimal and fleeting attention before making a decision.

We study one end of this spectrum by analyzing data from an eBay and MissionFish (a partner of eBay) randomized control trial testing nine scripts that solicited $1 donations at the point of checkout for individuals purchasing an item on the eBay auction shopping website. Individuals also had an option to increase the gift beyond $1. As this is a fleeting decision made at checkout for only $1, we argue this is likely best categorized as a “System I” decision. Naturally, this cannot be perfectly categorized as such. For example, if someone has already deliberated extensively on which charities to support, they may simply react by remembering decisions already made in the past and repeating the decision. We consider this a critical contextual factor when interpreting the results of our study, in that we are starting by assuming that we are operating in System I decision-making mode and testing individuals' "intuitive" reactions to different donor appeals. An interesting extension would be to test the same scripts but in a more deliberative decision-making environment.

Several of the scripts focus on charity quality and effectiveness. Recent literature shows that individuals respond negatively to what they perceive as high overhead and administrative costs (Gneezy, Keenan, and Gneezy 2014), and that individuals are willing to pay more for information about overhead ratios than for information on claims of impact (Metzger and Günther 2016). This is not to say that quality signals have no effect, and indeed, some argue that one reason matching or leadership gifts work is through a quality signal mechanism (see Vesterlund 2003 for a theoretical analysis, and Karlan and List 2012 for experimental evidence). Furthermore, in work more directly relevant for the tests here, Karlan and Wood (2016) find that adding information about scientific evidence of impact to a direct marketing letter via postal mail to prior donors of an international poverty charity has no impact, on average, on giving. However, important heterogeneity was observed, in that larger prior donors responded positively to the information, and smaller prior donors responded negatively. This effect persisted even after controlling for income and education (aggregated at the zip code level, hence a far from perfect control for income and education). With appropriate caveats for the challenges in interpreting why some donors previously gave more or less, we posit that small prior donors may be behaving more as System I "intuitive" donors (i.e., not deliberating much about the donation), whereas the large prior donors are deliberating. As such, we would expect our results in our online experiment reported here to be more similar to the small prior donors in Karlan and Wood (2016) and to potentially respond negatively, relative to other appeals.

Methods

Study Population

Our sample frame consists of eBay users who made a purchase on the American eBay site, www.ebay.com, in one of three weeks beginning January 9, 2011, January 23, 2011, or February 27, 2011. There were no restrictions as far as we know on being included in the study, and so all individuals who made a purchase on the site during the weeks the intervention was running participated in the study.

Study Design

A message appeared on the full sample frame at the confirmation step in the payment process. Figure 1 shows an example of the display individuals saw on the eBay website. Individuals were given the option to make a donation, defaulted to be $1, to their payment in support of the charity mentioned in the script. Participants in a given week were randomly assigned without any stratification to receive one of 22 messages (which consists of permutations of nine different content messages and three different charities). Each of the three weeks differed (non-randomly) in terms of the set of messages over which eBay randomized. Analysis will control for the week of the transaction. Table 1 provides summary statistics for each of the messages and charities. Aside from the differences in message provided, the display for each individual was identical. The donation was directly added to the bill presented by eBay and could be cleared with the rest of the amount due for the transaction.

Study Intervention

Individuals were shown one of 22 one-line messages at the point of checkout on eBay. These scripts varied along two dimensions: nine different content messages (popularity, fiscal efficiency, impact per dollar donated, impact per dollar donated with reference to scientific evidence for the specific program, scientific evidence for the specific program, scientific approach used at the organization, scientific approach used at the organization and matching grant, expert signal by naming Hewlett Foundation as a supporter, and expert signal without naming any particular expert), and three different charities (Pratham, Innovations for Poverty Action, and UNICEF).

Table 2 presents the specifics of each of the scripts and also identifies how we categorized each into attribute qualities for the sake of carrying out a regression to examine how attributes influence likelihood and amount of donation.

We selected three different nonprofit organizations: two less well-known charities (Innovations for Poverty Action and Pratham), and one very well known multilateral fund (UNICEF). Pratham is based in India and focuses on childhood education, and Innovations for Poverty Action is a research and policy organization headquartered in the United States.[2] The specific Innovations of Poverty Action program mentioned in the messages related to child health in Kenya. UNICEF is a widely known multilateral organization targeting child well-being globally. Including multiple charities in the study allows us to ensure that sentiments toward particular organizations can be controlled for when detecting the impact of the different messages.

Certain limitations on the information available for UNICEF meant that the study was set up to run four of the treatments with all three organizations and the remainder with only Pratham and Innovations for Poverty Action. During the study, a technical error related to the display of messages in the eBay platform resulted in only four different messages being displayed for Innovation to Poverty Action: the scripts relating to expert signal, matching grant, scientific evidence for the specific program, and expert signal by naming Hewlett Foundation as a supporter were missed and thus data for donations related to this content for Innovations for Poverty Action cannot be included in the analysis. All other messages were deployed as expected.

Randomization

Individuals were randomized at point of payment through eBay's internal website programming, and the principal investigators were not privy to the specific algorithm used to randomize the messages. A calendar of messages by organization and by week was prepopulated and sent to the client before the start of the study.

Study Outcomes

The two outcomes in this study are whether the individual made any donation, and the average amount given for each treatment in each week (note that we do not have the individual-level data on the size of each donation, just the average for the treatment cell by week). No other data are available. Note that in the tables, the percentage of donations is given in tenths of basis points, so the value of the binary donated outcome is either 0 or 1000, not 0 or 1.

Sample Size and Statistical Analysis

We observed 38,927,073 eBay purchases. We do not know how many of those are multiple purchases by a single user. The randomization was done by transaction, and thus if a user bought more than one item, they were rerandomized for each transaction, independently of their last treatment assignment.

We conducted two sets of analyses. Both employ ordinary least squares (OLS) with one of two dependent variables (a binary for “donated anything,” and the average amount donated).

The key independent variables in the first specification are indicator variables for eight of the nine content treatments. The specification also includes controls for the week of the experiment and the charity (because the randomization was conditional on week and charity). These results are presented graphically in Figure 2. Point estimates and standard errors for each treatment group are provided in the comments of the figure.

The key independent variables in the second specification are attributes of the content treatments. We assigned all treatments to a set of six attributes, since some of the messages overlap in the underlying theory they are intending to capture. The six attributes are as follows: message length, depiction of charitable activity, quantification of impact, matched funds, scientific evidence, and expert signal. Table 2 shows the mapping of the specific messages to these attributes. Table 3 presents the OLS regression results, examining how each attribute predicts likelihood of donating and average donation size. These specifications, as with the first specification, include control for charity and week.

Results

Figure 2 presents the main results comparing the proportion of individuals who donate in response to each of the nine scripts (after controlling for charity and week). Given the sample size, the confidence intervals are small, and for almost all pairwise treatment comparisons, we can reject a null hypothesis of equality. Table 1 presents the means for each treatment, broken down by charity, and reports the proportion who give (thus, it presents results that are similar to those shown in Figure 2, except without controls for week) and the average amount donated.

Table 3 presents what we consider the main results, testing the impact of each attribute. The omitted category is the "popularity" treatment, which is coded as zero for all attributes. Thus all results in this table are the effect of a particular attribute compared to the popularity treatment. For the linear probability model (column 1 in the table) to predict likelihood of giving, we find, in order of magnitude, the following point estimates: matching (1.069, standard error [se] = 0.077), depiction of charitable activity (0.886, se = 0.030), quantification of impact (0.390, se = 0.039), scientific evidence (0.107, se = 0.016), and expert signal (0.027, se = 0.037). The coefficient on number of words in the message is –0.062 (se = 0.009), which means that the effect of going from the longest to the shortest message generates the same treatment effect as the quantification of impact treatment (relative to the popularity message, which is the omitted variable).

Column 3 reports the treatment effects on the average amount given. Matching funds generates the largest treatment effect. The main change in ordering, compared to column 1, is for scientific evidence, which lowers average amount given compared to the omitted category (popularity). In contrast, for likelihood of giving, the scientific evidence generated a small (relative to the other treatments) but positive treatment effect. Furthermore, the expert signal did not generate a statistically significant treatment effect on likelihood of giving, but it did lead to a statistically significant increase in average amount given.

Conclusion

To interpret our results, we start by assuming that the decision-making environment triggered System I “intuitive” thinking. We then use this experiment to learn which treatments work well in a no-deliberation, “intuitive” decision-making environment. The results are, ahem, fairly intuitive:

- shorter messages are good;
- matching grants (which is a common marketing tool and hence requires little thought) work well;
- depiction of charitable activities works well (it provides immediate and tangible understanding of what an organization does);
- quantification of impact does not work as well (this requires thinking: is $1 for 2 years of medicine a good deal? Is this a credible deal?);
- popularity does not work well; and
- scientific evidence has a weak result on the likelihood of giving and a negative result on average amount given.

On a practical level, there are many retail sites, both in person and online, which promote giving. These results are likely relevant for such efforts.

We stress obvious caveats: mapping these scripts to specific theories is difficult and tenuous. Furthermore, we lack any further data on the donors, which could be used to test richer theories. Further research examining heterogeneity across donors would be fruitful. In addition, we believe it would be fruitful to test the efficacy of these types of treatments in an environment that allowed researchers to randomize deliberation. To do so would allow us to make stronger statements than we can from our current data about what intuitive versus deliberative individuals respond most to for charitable giving. In addition, it would inform us about modeling of charitable giving more generally.

References

Deaton, A. 2009. "Instruments of Development: Randomization in the Tropics, and the Search for the Elusive Keys to Economic Development." The Keynes Lecture, British Academy, London, January 2009.

Gneezy, U., E. A. Keenan, and A. Gneezy. 2014. "Avoiding Overhead Aversion in Charity." *Science* 346 (6209): 632–635. doi: 10.1126/science.1253932.

Kahneman, D. 2003. "Maps of Bounded Rationality: Psychology for Behavioral Economics." *American Economic Review* 93 (5): 1449–1475.

Karlan, Dean, and John List. 2012. "How Can Bill and Melinda Gates Increase Other People's Donations to Fund Public Goods?" Cambridge, MA: National Bureau of Economic Research. doi: 10.3386/w17954.

Karlan, Dean, and Daniel H. Wood. 2016. "The Effect of Effectiveness: Donor Response to Aid Effectiveness in a Direct Mail Fundraising Experiment." *Journal of Behavioral and Experimental Economics* 66: 1–8. doi: 10.1016/j.socec.2016.05.005.

Metzger, Laura, and Isabel Günther. 2016. "Making an Impact? The Relevance of Information on Aid Effectiveness for Charitable Giving. A Laboratory Experiment." Working paper, ETH Zurich, Zurich, Switzerland.

Vesterlund, Lise. 2003. “The Informational Value of Sequential Fundraising.” *Journal of Public Economics* 87 (3): 627–657.

Figure 1 Screenshot for donation page on eBay website.

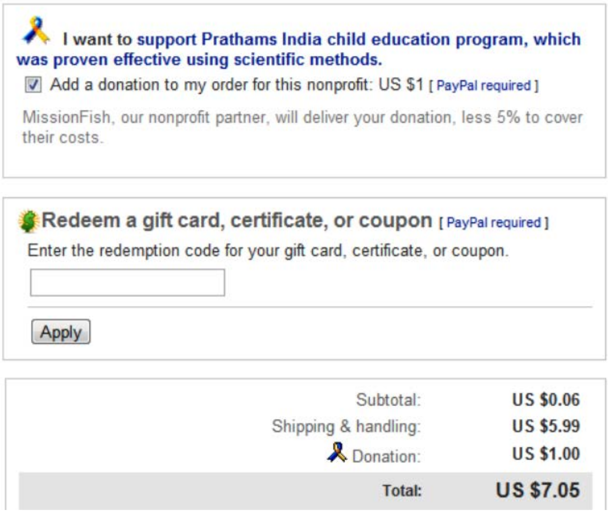
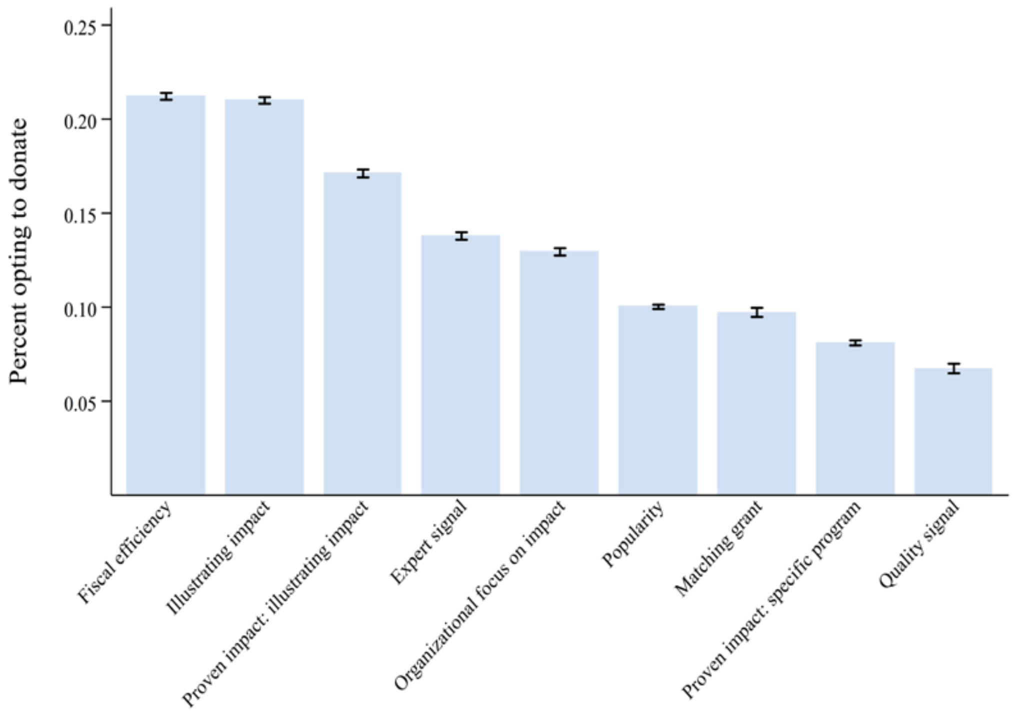


Figure 2 Donation rates by treatment



Note to Figure 2: Point estimates and standard errors (in brackets) by treatment in terms of tenths of basis points are as follows: Fiscal Efficiency (Omitted Category): 2.872 (0.036), Illustrating impact: -0.079 (0.025), Proven impact: illustrating impact: -0.137 (0.027), Expert signal: -1.324 (0.032), Organizational focus on impact: -0.201 (0.027), Popularity: -1.277 (0.022), Matching grant: -0.240 (0.035), Proven impact: specific program: -0.544 (0.023), Quality signal: -0.539 (0.036).

Table 1 Summary statistics.

Treatment	1000 × Likelihood of Donating $1 (average amount donated, $) IPA	Pratham	UNICEF	All
Treatment 1: Fiscal efficiency	1.584 (1.295)	0.811 (1.285)	4.599 (1.274)	2.121 (1.284)
	n = 1,444,940	*n* = 3,028,671	*n* = 1,913,616	*n* = 6,387,227
Treatment 2: Illustrating impact	3.586 (1.278)	1.499 (1.301)	1.953 (1.282)	2.099 (1.290)
	n = 1,565,669	*n* = 3,361,685	*n* = 2,107,862	*n* = 7,035,216
Treatment 3: Proven impact: illustrating impact	2.660 (1.265)	1.017 (1.265)		1.711 (1.265)
	n = 1,634,365	*n* = 2,235,467		*n* = 3,869,832
Treatment 4: Expert signal		0.375 (1.335)	1.586 (1.317)	1.378 (1.320)
		n = 596,859	*n* = 2,881,419	*n* = 3,478,278
Treatment 5: Organizational focus on impact	1.551 (1.173)	1.254 (1.225)	-	1.294 (1.218)
	n = 457,046	*n* = 2,936,690		*n* = 3,393,736
Treatment 6: Popularity	0.959 (1.271)	0.593 (1.211)	1.471 (1.297)	1.002 (1.259)
	n = 2,297,672	*n* = 2,675,654	*n* = 2,534,642	*n* = 7,507,968
Treatment 7: Matching grant		0.972 (1.307)	-	0.972 (1.307)
		n = 1,650,515		*n* = 1,650,515
Treatment 8: Proven impact: specific program		0.810 (1.235)	-	0.810 (1.235)
		n = 4,553,399		*n* = 4,553,399
Treatment 9: Quality signal		0.674 (1.274)	-	0.674 (1.274)
		n = 1,050,902		*n* = 1,050,902
All	2.049 (1.270)	0.963 (1.261)	2.248 (1.295)	1.481 (1.271)
	n = 7,399,692	*n* = 22,089,842	*n* = 9,437,539	*n* = 38,927,073

Notes: IPA, Innovations for Poverty Action. Each cell reports the proportion of views of each script for each charity that generated a donation (reported in tenths of basis points), the average amount donated of those that donated (in parentheses), and the sample size *n* per cell. The exact message scripts for each treatment are provided in Table 2.

Table 2 Treatment scripts and assigned attributes.

Treatment	Text of Script	Charity	Message Length (number of words)[a]	Depiction of Charitable Activity	Quantification of Impact	Matched Funds	Scientific Evidence	Expert Signal
Treatment 1: Fiscal efficiency	I want to support [organization's program], which has low overhead expenses.	All	14.7	X				
Treatment 2: Illustrating impact	I want to support [organization's program]. $1 provides [recipient] with one [program relevant outcome].[b]	All	19.0	X	X			
Treatment 3: Proven impact: illustrating impact	I want to support [organization's program], proven effective with scientific methods. $1 provides [recipient] with one [program relevant outcome].[b]	IPA, Pratham	19.0	X	X		X	
Treatment 4: Expert signal	I want to support [organization], whose methods have been approved by experts in international development.	Pratham, UNICEF	15.0					X
Treatment 5: Organizational focus on impact	I want to support [organization], which uses scientific methods to fight poverty.	IPA, Pratham	12.0				X	
Treatment 6: Popularity	I want to support [organization], one of the top nonprofits on eBay.	All	12.0					
Treatment 7: Matching grant	I want to support [organization], which uses scientific methods to fight	Pratham	21.0			X	X	

	poverty. My gift will be matched by a major foundation.					
Treatment 8: Proven impact: specific program	I want to support [organization’s program], which was proven effective using scientific methods.	Pratham	16.0	X	X	
Treatment 9: Quality signal	I want to support [organization], whose anti-poverty programs have been evaluated and supported by the Hewlett Foundation.	Pratham	17.0			X

Note: IPA, Innovations for Poverty Action.

[a] Average across organizations. Actual length differs slightly by organization.

[b] For program-relevant outcomes, the content was as follows. For Pratham, an Indian child education program, for which $1 provides a child with one semester of education; for IPA, a Kenya child deworming program, for which $1 provides a child with 2 years of medicine. For UNICEF, a safe drinking water for kids program, for which $1 provides a child with 40 days of clean water.

Table 3 Effect of message attribute on likelihood of donating: OLS, probit.

Attribute	Donated (1000/0): OLS	Donated (1000/0): Probit	Average amount donated ($/1000): OLS
Message length (number of words)	-0.062	-0.047	-18.743
	(0.009)***	(0.002)***	(0.006)***
Depiction of charitable activity	0.886	0.300	81.583
	(0.030)***	(0.008)***	(0.019)***
Quantification of impact	0.390	0.211	91.210
	(0.039)***	(0.009)***	(0.028)***
Matched funds	1.069	0.545	231.851
	(0.077)***	(0.021)***	(0.039)***
Scientific evidence	0.107	0.068	-22.644
	(0.016)***	(0.004)***	(0.012)***
Expert signal	0.027	0.094	99.567
	(0.037)	(0.007)***	(0.019)***
Organization 1 (IPA)	-0.545	-0.164	-30.558
	(0.025)***	(0.005)***	(0.015)***
Organization 2 (Pratham)	-1.604	-0.381	-34.075
	(0.021)***	(0.005)***	(0.015)***
Week 1	-0.011	0.020	-12.001
	(0.022)	(0.005)***	(0.022)***
Week 2	0.185	0.058	-40.197
	(0.020)***	(0.005)***	(0.021)***
Constant	2.697	-2.348	1525.007
	(0.121)***	(0.029)***	(0.092)***
Number of observations	38,927,073	38,927,073	38,927,073
Mean of dependent variable	1.481	1.481	1270.8

Notes: IPA, Innovations for Poverty Action; OLS, ordinary least squares. Estimates for columns 1 and 2 are in tenths of basis points (i.e., the dependent variable is either 1000 or 0, and the independent variables are indicator variables equal to 1 or 0). Depiction of charitable activity corresponds to treatments 1, 2, 3, and 8 in table 1; quantification of impact corresponds to treatments 2 and 3; matched funds corresponds to treatment 7; scientific evidence corresponds to treatments 3, 5, 7 and 8 expert signal corresponds to treatments 4 and 9. Probit results are marginal effects. Robust standard errors are shown in parentheses. The symbol *** indicates significance at the 1 percent level.

Notes

1. We thank MissionFish, eBay, and Clam Lorenz for collaboration and making the data available, and Matthew Grant and Nicole Mauriello at Innovations for Poverty Action for management and data analysis. This manuscript was prepared as a chapter for *The Economics of Philanthropy*. We also thank the book's editors Kimberley Scharf and Mirco Tonin. All opinions and errors are our own.

2. Disclosure: Karlan is founder and chairman of Innovations for Poverty Action, and at the time of this experiment was the executive director.