Giving is Caring:
Understanding Donation Behavior through Email

Yelena Mejova
Yahoo Labs
Av. Diagonal 177
E-08018 Barcelona, Spain
ymejova@yahoo-inc.com

Ingmar Weber and Venkata Rama Kiran Garimella
Qatar Computing Research Institute
Doha, Qatar
{iweber,vgarimella}@qf.org.qa

Michael C Dougal
University of California, Berkeley
210 Barrows Hall
Berkeley, California 94720
mdougal@berkeley.edu

ABSTRACT

Every day, thousands of people make donations to humanitarian, political, environmental, and other causes, a large amount of which occur on the Internet. The solicitations for support, the acknowledgment of a donation and the discussion of corresponding issues are often conducted via email, leaving a record of these social phenomena. In this paper, we describe a comprehensive large-scale data-driven study of donation behavior. We analyze a two-month anonymized email log from several perspectives motivated by past studies on charitable giving: (i) demographics, (ii) user interest, (iii) external time-related factors and (iv) social network influence. We show that email captures the demographic peculiarities of different interest groups, for instance, predicting demographic distributions found in US 2012 Presidential Election exit polls. Furthermore, we find that people respond to major national events, as well as to solicitations with special promotions, and that social connections are the most important factor in predicting donation behavior. Specifically, we identify trends not only for individual charities and campaigns, but also for high-level categories such as political campaigns, medical illnesses, and humanitarian relief. Thus, we show the extent to which large-scale email datasets reveal human donation behavior, and explore the limitations of such analysis.

Author Keywords

donations; email; charities; social factors

ACM Classification Keywords

K.4 Computers and Society: Miscellaneous; J.4 Social and Behavioral Sciences: Sociology

General Terms

Human Factors

INTRODUCTION

The development of the Internet as a marketplace has brought about an era of online giving. The largest source of charitable giving are individuals, donating $211.77 billion in 2010[1]. The 2008 US presidential campaign saw a new wave of online mobilization, with websites like *moveon.org* collecting millions of dollars for political campaigns[2]. Campaigns used "money bombs" to set single day fundraising records. On December 16, 2007 a fundraising effort raised $6 million for the presidential candidate Ron Paul, largely by having a strong online presence[3]. Charities are increasingly fundraising online, and understanding this market may mean thousands of additional donations.

While past literature has cited the potential of the Internet for fundraising and other charitable activities [18, 21], relatively few studies look at a web-based dataset on actual donor behavior. Commonly, such studies perform user surveys focusing on a single website [9]. However, the number and diversity of potential donors on the web, as well as organizations seeking their support, provides an unprecedented opportunity for a data-driven study of the factors contributing to an individual's donating behavior online.

We take inspiration for this study from the copious sociology and psychology literature on individuals' motivations for charitable giving. Social pressure and reward have been found to be powerful motivators [8], as well as social identity in the form of social norms such as gender and income [2]. However, in a recent study of online charitable behavior on GiveIndia, Shier & Handy [41] find an inverse relationship between perceived social pressure to donate (both explicit like solicitations and implicit like other members of social group donating) and likelihood of donating to a cause. This finding puts in question the role of social propensity in an individual's motivations. Guided by these findings, we model individuals' charitable motivations using both internal and external factors – all extracted from a large set of web data.

---

[1] http://www.nps.gov/partnerships/fundraising_individuals_statistics.htm
[2] http://www.factcheck.org/2010/08/moveonorg/
[3] http://www.npr.org/templates/story/story.php?storyId=17332414

Specifically, this paper contributes a unique multi-faceted analysis of donors' behaviors and interaction with the world as captured by four respects: (1) demographics, (2) topical interest, (3) social influence, and (4) external influence. We examine the communication of hundreds of thousands of anonymous Yahoo! users with a wide range of organizations, including political campaigns, humanitarian efforts, and special-interest groups. Interactions with organizations are mined from a collection of anonymized emails spanning a two-month period of July 19 - September 19, 2012, including interactions with 480 organizations. This analysis is further enriched by demographic information about the users, their interactions with the emails, and an induced social network. Thus, we were able to explore a variety of factors contributing to donations. Finally, throughout this analysis we discuss the benefits and shortcomings of using email for human behavior research. We conclude this study with a set of guidelines to organizations aiming to solicit contributions online, and especially via email.

RELATED WORK

The literature identifies several main factors that influence whether or not an individual will donate to a particular charity: individual capacity and general inclination to donate, individual's interest in the charitable group, social influence and pressure, and the institutional ability of the charity to solicit donations. We briefly discuss these here.

Individual capacity and inclination.   Individual capacity generally refers to "a person's human and financial resources (level of education and income, respectively)" [41]. To the extent that charities provide collective goods, they face a collective action problem wherein the rational self-interested individual prefers to free-ride on other people's donations rather than make a donation herself [32]. Individual inclination "typically refers to the social and psychological propensity that a person has towards giving" [41]. The degree to which an individual is altruistic, the extent to which a donation (or lack of thereof) will influence his personal reputation, the psychological benefits that accrue from making a donation, and his values are all important forces that drive charitable giving [8]. More generally, donating can provide a psychological benefit by reinforcing an individual's positive self-image [8] and providing them with an "empathetic joy" [5] or "warm glow" [3]. Related to the notion of self-image, Beckers observes that the Name-Letter Effect [31], where people have a general preference for things sharing their first name initial, can be used to increase donations [6]. For example, a Janette receiving a phone call from a Jane would have a higher donation probability. In this study, we model internal and behavioral user attributes using demographics and emailing behavior.

Interest.   Individuals give to organizations because they care about their charitable work. Bekkers [7] finds that people are more willing to contribute to charities that are "more needy and deserving beneficiaries", even if these beneficiaries are geographically far away. Also, people tend to donate more to organizations when they perceive them to have a higher measure of performance and accountability [37]. Unlike in previous survey-based studies which rely on self-reporting, in our data-driven study we measure the interests of users using their daily interactions with friends and the outside world.

Social influence.   Donors can also be thought of as "impure altruists" to the extent that their behavior is driven by the social and psychological benefits of donating [4, 8]. In the context of a tightly knit group, such as a small church congregation, donation behavior can be monitored to socially reward donors and punish non-donors [8]. In contrast, online donations and solicitations for charitable donations made via email are made in relative anonymity. Nevertheless, social identity can still play a large role in anonymous online donations. Social identity theory tells us that positive in-group identity is important to individual self-esteem [43] and that making a donation can reinforce that identity in response to out-group threat [13]. Choosing donations made in one's name over being paid has also been shown to increase one's credibility in the context of user-generated content [24]. In comparison to in-person solicitation, email solicitation can "increase the salience of social identity" related to a charitable group while reducing "cognitive perceptions of interpersonal differences" within a group [11]. Such social interactions can be glimpsed in the email use of the users in our dataset.

External influence.   Finally, the ability of an organization to solicit donations is an important institutional factor in charitable giving. Charities can provide selective incentives to help overcome collective action problems [32], such as giving donors access to exclusive events. Donors are also responsive to decreases in the cost of donating, such as tax incentives [33] and offers to match contributions [25]. We explore both the nature of successful solicitations, as well as the effect of major events on donation behavior.

While there is a vast social science literature on charitable donations, there are relatively few studies that examine either online charitable donations or email-based charitable donation solicitations. Recently, studies of large donation websites such as *www.give-india.com* [41] and *www.kiva.org* [27] have allowed for a large population of study. Specifically, [27] build a text classifier that identifies reasons for why people donate to charities from their free-text elaborations on the topic. In contrast, our study provides a more involved analysis of the donors' interactions with the charitable organizations, as well as their social network and the rest of the world.

A number of researchers have looked at the role that new technology plays concerning fundraising. Goecks et al. [17] study 150 instances of technology for nonprofit fundraising, including Facebook Causes and Network for Good. They claim that technology like email helps in making visible an immediate need and framing a contribution as helping individuals rather than groups. Technology can also provide information about what other people have donated and legitimize small contributions. Hayes et al. [22] present several studies related to "modern" donation. They stipulate that "philanthropic organizations wishing to connect to donors strive to become part of their everyday 'small world' ". Using the American Red Cross website as an example, they illustrate how interactivity and social networking tools have dras-

tically increased the number and depth of connections that can exist between donors and causes. Popups asking for immediate contributions are found to be ineffective in [47]. The same authors find evidence though that they can help to remind users to contribute, leading to a 23% increase in contributions. However, this effect wanes as users habituate to the popups. After the 2010 earthquake, special SMS donation numbers were setup leading to over $7M in micro donations. The technology-enabled ease of the system - just sending and SMS to a dedicated number - is widely regarded as crucial in this success [34]. Going beyond donations, Rotman et al. [35] present a discussion of open research questions related to social media and their use for engaging people in activism for certain causes.

Recently, researchers have looked at *crowdfunding* platforms to understand the arguably altruistic motivation of the contributors. Muller et al. [29] studied an intra-organization crowdfunding system where participants were each given USD 100 by the organization to "use or lose". Considered a successful pilot, it showed the importance of the social context and of achieving real change in their workplace. Though interesting, their findings might not be transferable to donations to "external" charities. Greenberg et al. [19] go beyond the focus on financial committments and apply Resource Exchange Theory to model crowdsourcing. They find that platforms support the implicit exchange of all six resources (money, love, information, status, goods, and services). The "tragedy of the commons" is a classic problem in game theory where each player's individually optimal move is not to contribute to a common good. Sakamoto et al. [36] try to tackle this problem through crowdfunding. In their envisioned system, people contribute "ageing money" to projects in their direct environment to keep them sustainable. Projects range from public toilets to office areas. They built a first prototype system for mobile devices but have not yet provided any evaluation of real results. Knowing that your crowdfunding application will or will not be successful could help you to improve your submission. Greenberg et al. [20] study the corresponding binary classification problem by applying machine learning classifiers to learn the concept of a successful online crowdfunding project at the time of project launch. The features used include the set monetary goal in USD, the topic category, how popular the person in chare is on Twitter and other features. Ultimately, crowdfunding can be viewed as a "creativity support tool" and Kuoger et al. [26] apply the "Genex (Generator of Excellence) Framework" to model platforms. They link the four phases from Genex, namely, Collect, Relate, Create and Donate to features in crowdfinancing platforms.

OBSERVING DONATIONS IN EMAIL

Data Gathering
We begin by describing our anonymized email dataset, which spans a two-month period of July 19 – September 19, 2012.[4] In the following, a "user" refers to a Yahoo! email user. For

---

[4]Data originated from users who voluntarily opted in for their mail to be analyzed for research purposes. **http://info.yahoo.com/ privacy/us/yahoo/**

Table 1. Top 10 charity domains by donation volume

barackobama.com	US President Barack Obama's Campaign
mittromney.com	Mitt Romney for President Campaign
cancer.org	American Cancer Society
lls.org	Leukemia & Lymphoma Society
alz.org	Alzheimer's Association
nmss.org	Multiple Sclerosis Society
worldvision.org	World Vision (humanitarian)
redcross.org	American Red Cross
irusa.org	Islamic Relief USA
stjude.org	St. Jude Children's Research Hospital

our analysis, all user identifiers (= email addresses) were replaced by random numbers ("hashes"). For simplicity, we will call any entity that may seek a donation a *charity*, regardless of its actual charitable status, thus generalizing the word to refer to political campaigns, food drives, human rights groups, etc.

An initial set of charities was gathered by scraping the lists of top charities posted on reputable websites like Forbes[5] and USNews[6]. The top 100 US political campaign organizations by money raised to date in the 2012 election cycle were also added to the dataset[7]. Finally, we compiled a list of prominent news stories in that time span, using Wikipedia Current Events portal[8] and compiled lists of charities which may be affected by them. This resulted in 480 internet domains each belonging to a charity.

In order to identify an email from a charity, we matched the domain (or subdomain) of the URL in the *from* field of the email to the list of our charities. Furthermore, we identified emails thanking the user for a donation using a manually compiled set of regular expression rules. We manually checked the accuracy of these rules (having an average 86% accuracy). Table 1 shows the top 10 domains which received the most donations in our dataset. They include political campaigns, humanitarian organizations, and medical illness associations.

We then identified three groups of users who may be of interest. The *donors* are the users who have received an email from any of our 480 charities that thanks them for a donation. This set contains on the order of 100,000 users. The *interested* users are those who have received an email from one of the charities, but who have not donated, or rather received a "thank you" email. We down-sample these users to get a set of roughly the same size as the previous. Finally, we compare these groups to a *general* random subset of 10,000 users who have received at least one email in that time period and who are not in the above samples. Incoming and outgoing emails were gathered for all of these users, and the data anonymized. In total, around a billion incoming and several million outgoing emails were collected. These included metadata, such as (anonymized) *from* and *to* fields, the date

---

[5]**http://www.forbes.com/lists/2011/14/ 200-largest-us-charities-11_rank.html**

[6]**http://www.usnews.com/usnews/biztech/charities/ lists/intl_deve-lopment.htm**

[7]**http://www.fec.gov/data/CommitteeSummary.do? format=html&election_yr=2012**

[8]**http://en.wikipedia.org/wiki/Portal:Current_ events**

email was sent/received, and the subject. The body of the emails was not collected.

In order to aggregate over the various charities, we selected the charities that received at least 100 donations within these two months and annotated them for topical category. The 49 charities in this list accounted for 99.3% of all donations in our dataset. The ten categories are (in the descending order of prominence) *Medical*, *Humanitarian*, *Politics*, *Environmental*, *Religious*, *US Military*, *Children*, *Public Broadcasting*, *Animals*, and *Internet*. Additionally, a set of distinguishing keywords were manually compiled such that if one is found in a subject of an email, it is most likely about the given topic.

Spam Filtering

As the incoming mail comes with a *spam* flag, we investigate the distribution of this flag in the emails sent by the charities. According to this flag, the email may be directed to the inbox, or to bulk or spam folders. We find that there is a significant number of emails from these charities that do not end up in the users' inbox. Although most organizations, such as *cancer.org*, *lls.org*, and *redcross.org* have less than 10% of their emails flagged as spam, more than half of the emails from *stjude.org* (St. Jude Children's Hospital) and *dscc.org* (Democratic Senatorial Campaign Committee) are labeled as bulk and do not reach the user's inbox.

We further investigate this by gathering user action statistics. These include instances of users reading, deleting, and forwarding emails, as well as tagging them as *spam*. We find that the proportion of emails tagged as spam to those read is the greatest for *wikimedia.org* – an organization responsible for Wikipedia and other projects – at 0.55. Indeed, upon examination we find the domain may have been "hijacked" by spammers. Thus, the server-side spam-filtering algorithms which are out of the user's control may affect the user's perception of the charity, and we discuss the possible effects the spam filter may have on donation behavior in the Discussion section. Also, when considering the user's interaction with the charity, we exclude those emails tagged as spam.

DONATION BEHAVIOR ANALYSIS

Demographics

Sociological studies on motivation often link motivation to group membership and identity definition. Demographics such as gender, age, and income, play an important role in an individual's formation of their self-image [39]. In order to determine if users' demographic characteristics play a role in their donation behavior, we collect the self-reported demographic information from the user profiles. This information includes age, gender, and zip code. In the case the user's zip code is in US, we use census data to obtain localized statistics such as the percentage of bachelors degrees and median household income[9]. We aggregate these for the users donating to charities within the ten topics. The age, gender (male = 1), percentage of bachelor degrees and median household income are shown in Figure 1 (note, the 95% confidence intervals are too small to plot).

---

[9]**http://www.census.gov**

Table 2. US Presidential Election exit polls

Age	Obama	Romney	% total
18-29	60	37	19
30-44	52	45	27
45-64	47	51	38
65 or over	44	56	16
Gender	Obama	Romney	% total
Male	45	52	47
Female	55	44	53
Bach Degree	Obama	Romney	% total
College graduate	50	48	47
No college degree	51	47	53
Family Income	Obama	Romney	% total
Under $50,000	60	38	41
$50,000 - $99,999	46	52	31
$100,000 or more	44	54	28

Several interesting trends become evident. First, note the average age is usually between 40 and 55 (even after removing the small number of instances where users claim age of > 100), showing that donors in our dataset tend to be of an older generation. Donors to political and military causes, as well as donors to public broadcasting, tend to be older. Environmental and internet (including sites like *wikipeda.org*) causes tend to attract slightly more male donors, whereas donors to animal-related causes tend to overwhelmingly be female. Public broadcasting, which is mostly operated by non-profits in US, attracts donors from neighborhoods with higher household income and a higher percentage of bachelor degrees.

Figure 2 shows the same distributions for two major political campaigns: *barackobama.com* and *mittromney.com*. On average, Romney donors are older and live in more affluent areas[10], while Barack Obama attracts significantly more women. When we compare online donation behavior to the US Presidential Election exit polls[11], we find that donors exhibit a similar gender gap as voters, but that Barack Obama's voters are much younger than his donors. This is consistent with the finding of [40] that political donors are generally older and more affluent than the electorate as a whole. Note that we show the original statistics instead of computing an "average voter" due to vagueness in classes like "65 or over".

Donor Interest

Next, we attempt to measure the effect of the user's interest on donation behavior. Intuitively, people interested in a cause are more likely to donate. First, we note that for most part donors are more likely to receive emails from the charity than those who do not donate but still receive some emails (the "interested" group). For example, Figure 3 shows the distribution of emails received from *cancer.org* and *redcross.org* by users in the two groups. We see that even though both charities deal with health issues, the difference between the

---

[10]Remember that we do not have data on individual income and education, only zip code level information. The average donor to any campaign is likely to be better educated and have a higher household income than their neighbors [40].
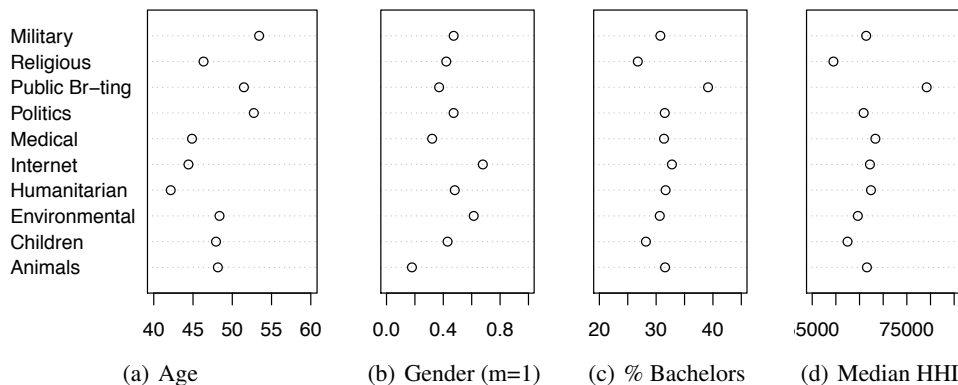
[11]**http://elections.msnbc.msn.com/ns/politics/2012/all/president/#exitPoll**

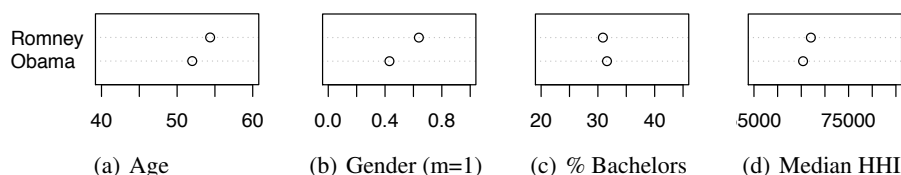**Figure 1. Demographics of donors to causes in 10 topical categories**

Figure 2. Demographics of donors to two political campaigns

number of received emails between the two groups of users is more pronounced for the former than the latter.

However, just receiving emails from a charity is not a sure sign of interest. Instead, we estimate interest by classifying the subjects of the incoming and outgoing emails according to the ten topical categories. Specifically, we use regular expressions with manually-compiled keywords such as one below for *Politics*:

$(politic\backslash w^* \mid republican\backslash w^* \mid democrat\backslash w^* \mid$

$election\backslash w^* \mid elect\backslash w^* \mid tax\backslash w^* \mid obama\backslash w^* \mid$

$romney\backslash w^* \mid white\ house \mid senat\backslash w^*)$

Upon manual evaluation of a selection of 50 email titles for each topic, we find the average accuracy of these regular expressions to be 86.2% (although coverage would be more difficult to check). Figure 4 shows the topical shares of incoming and outgoing emails for the *donors*, users who have received an email from a charity but have not donated (*interested*), and a general subset of users (*general*) with 95% confidence intervals. Note here that donors and interested are not all of the people in our *donors* or *interested* groups, but people who have donated to (in case of donors) or received an email from (interested) at least once to the cause in the category.

Considering the incoming emails, we can see that in most cases (8 out of 10), the donors are more likely to receive emails about such topics than other groups. Similarly, users whose (outgoing) emails mention *animals*, *politics*, or *religion* are drastically more likely to donate to these causes.
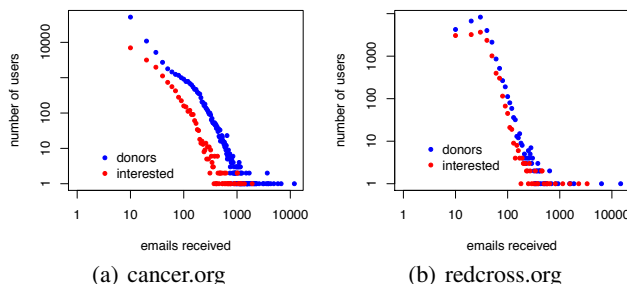

(a) cancer.org   (b) redcross.org

Figure 3. Number of users receiving an email from a charity in donor and interested groups

Most other topics do not show much difference between the behavior of the three groups. Note that we saw very few incoming and no outgoing emails about public broadcasting for all three groups. This is consistent with organizations like WGBH (Boston Public Broadcasting) and PBS (Public Broadcasting Service) raising money primarily through other means, such as on-air fundraising telethons. Similar observations can be drawn about *medical* charities and *humanitarian* charities, both of which have well-established organizations that promote donations offline. An alternative explanation is that animals, politics, and religion are more natural topics of conversation for animal lovers, political junkies, and the religiously devout than, say, public radio is for NPR listeners, such that our classifier is better able to pick up an interest in election outcomes than an interest in quality public radio. Nevertheless, the relationship between the content of incoming emails and donation may be indicative of the effectiveness of solicitation campaigns. We examine this next.
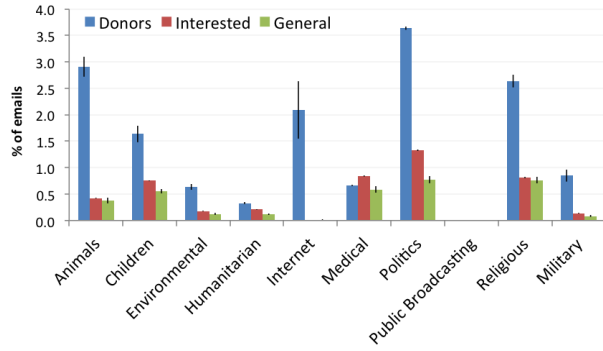
(a) Incoming emails (non-spam)
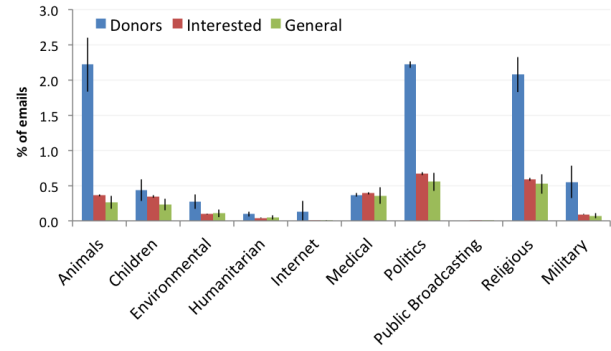
(b) Outgoing emails

Figure 4. Topical share of incoming and outgoing emails for the donors, interested, and general sets of users

Solicitation

Regardless of the user's interest in the charity, from the point of view of the charitable organization it would be useful to gauge whether donations can be encouraged by solicitation. We address this by looking at overall solicitation and donation volumes, as well as donation on a per-user basis.

Table 3. Cohen's kappa between donation volume and charity solicitations, top and bottom 10 charities

Domain	Category	kappa
toysfortots.org	Children	0.9379
firstfoodbank.org	Humanitarian	0.8070
tbn.org	Christian	0.7377
dav.org	US Military	0.7304
nationalmssociety.org	Medical	0.6873
earthisland.org	Environmental	0.6796
wikimedia.org	Internet	0.6659
smiletrain.org	Medical	0.6625
feedthechildren.org	Humanitarian	0.6423
stjude.org	Medical	0.6186
hsus.org	Animals	0.0223
cff.org	Medical	0.0216
arthritis.org	Medical	0.0205
pbs.org	Public Broadcasting	0.0046
diabetes.org	Medical	-0.0008
ucsusa.org	Environmental	-0.0379
autismspeaks.org	Medical	-0.0496
jdrf.org	Medical	-0.0670
barackobama.com	Politics	-0.0722
wgbh.org	Public Broadcasting	-0.0973

First, to understand the "burstiness" of the donation volume over time, we characterize the donations as happening in a "low", "medium", or "high" period. To do this, we compute the number of donations per day for each charity. We then order these days from least active days to most active, where activity is the number of donations, and then split these into the three terciles, each having a third of all of the days in the dataset.

Table 3 shows the Cohen's kappa between the (discretized) donation volume and incoming mail from that charity. We exclude from this analysis charities for which the donation activity was too sparse. Organizations high on this list rely more on fundraising emails or have a high precision: when

they send an email, they tend to get donations. For organizations towards the bottom the relationship is reversed: people donate on any day, regardless of whether a solicitation was sent on that day. In the full list of these organizations, 43% have Cohen's kappa greater than 0.3, suggesting a positive relationship between email solicitation and donations.

In order to determine whether solicitations are effective on personal level, we compute the probability that, given the user will donate to an organization, that he or she donates within some number of days of receiving a solicitation. We compare this probability to a baseline, that uses the total number of donations the user has made to this charity over the dataset. The results are shown in Figure 5. The probabilities were first aggregated over all users donating to a charity, and then over charities (resulting in a macro-average). The cumulative probability of a user donating after a solicitation is higher by approximately 8% than the baseline probability. The organizations which have the best response (donation) rate to their solicitations are largely concerning medical illness, children's charities, and humanitarian relief. The only political campaign among the top ten most successful solicitors list is dscc.com (Democratic Senatorial Campaign Committee), with a 15.4% difference between observed and simulated donation rates.

We note that both of these measures can only capture correlation rather than causation. If charities send out solicitations only in periods where a user would have donated anyway, then our measure will overestimate the effectiveness of solicitations. If, on the other hand, charities send out emails every single day then our measure would not detect any correlation, even though donations could still be caused by these constant solicitations. One of the challenges is determining the intent of a charity's email. For instance, a Granger Causality test informs us that it is the donations which predict other emails, since we find that the policy of following up on a donation on the next day is common in our dataset. To determine the causal effect of any given email or email campaign, charities should run randomized, controlled experiments that randomly assign prospective donors to treatment (an email or email campaign) and control (no emails) and measure the difference in donation behavior.
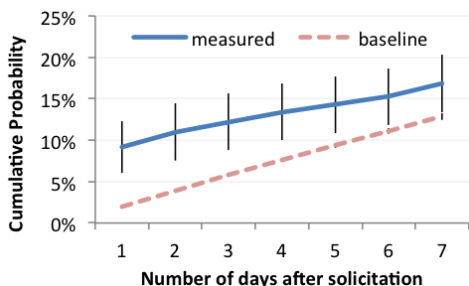
Figure 5. Cumulative probability (in %) of donation within n days of solicitation (with 95% confidence intervals)
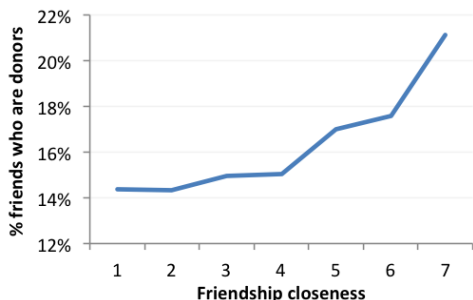

Figure 6. Percentage of donors' friends who are also donors (close friends = 7)

Social Network

Direct interaction between individuals has been shown to be important in information diffusion and has been used in marketing [23] and epidemiology [15]. Also, individuals tend to become friends with similar others, a phenomenon called *homophily* [28]. In order to examine the relationships between social ties and donation behavior, we extract a social network from the emails.

First, for each donor, we found all users they exchanged emails with, constraining these to users within the Yahoo! network. The average number of friends within this network was 47. Because our lists of donors and user who have received an email from a charity are complete, we can look up whether any of these friends are themselves in a *donors* or *interested* groups. In order to normalize the contribution of each user to our aggregate measures, we sample each user's contacts with replacement to get 100 friends. We measure the strength of the relationship between two users by the minimum of emails sent or received (that way one-sided relationships such as spam would not be considered a strong relationship). We then bucket the users by relationship strength from distant acquaintances to close friends (with each bucket containing the same number of users). Within these buckets, we calculate the percentage of donors (Figure 6). The spike in the seventh bucket, which contains the closest friends, is significant at a $p = 0.01$ level (compared to both buckets 1 and 6). When we compute these statistics on per-charity basis, we find similar results. Furthermore, we find that not only are friends of donors more likely to donate to that cause, but they are also more likely to receive solicitations from the same domain.

Not only do interests unite users, they also divide them. For instance, the two major opposing sides of the US political scene – one supporting Barack Obama, another Mitt Romney – are more connected within the party than across. In Figure 7 we show the percentage of users we found to have a connection (who have both sent and received emails to/from each other) over all users in the group. This divide has been previously observed both for the interlinkage network of political blogs [1] and for the social network of Twitter users [14]. While roughly 1 in 4 Americans have reported having regular contact with someone of the opposite political persuasion, almost none of our online donors have such a connection [30]. This extreme homophily is consistent with [30]'s finding that homophily increases political participation.

Figure 7. Percentage of donors having social ties with other donors within their own or opposing political groups

As we show in the Regression Analysis section, this relationship between friends is the most predictive factor of donation behavior when we consider interest, demographic, behavioral, and social variables.

External Influence

Next, we examine the role of external influences on donation behavior. A driving factor of charitable donations is awareness of need. In particular, donors' subjective perceptions of need is crucial [8][46]. Increased perception of need has been found to be positively related to university alumni support [48], volunteering in the community [44], and donations to international relief organizations [12]. The media can play an important role in facilitating this awareness. Increased media coverage of a particular earthquake, for example, tends to increase donations to that earthquake's relief efforts [42].

In order to capture the effect of current events on donation behavior, we manually examine the donation levels for select charities. Figure 8 shows the daily donation (in blue) and solicitation (in red) volumes. The first organization, *barackobama.com* is the official US Presidential campaign of Barack Obama, a Democrat and the second, *mittromney.com*, is that of his Republican opponent Mitt Romney. Here are several dates of interest:

July 31: "Photo going around on Facebook" (Obama)
Aug 11: Paul Ryan is Republican VP candidate
Aug 18: T-shirt promotion "I built this" (Romney)
Aug 27-30: Republican National Convention
Sep 4-6: Democratic National Convention

We see major spurs in donations following promotional emails like "Photo going around on Facebook" where a potential voter shows their support for Obama. These solicitations are not as effective on Romney's side, for example, when a promotion on August 18 did not result in a similar
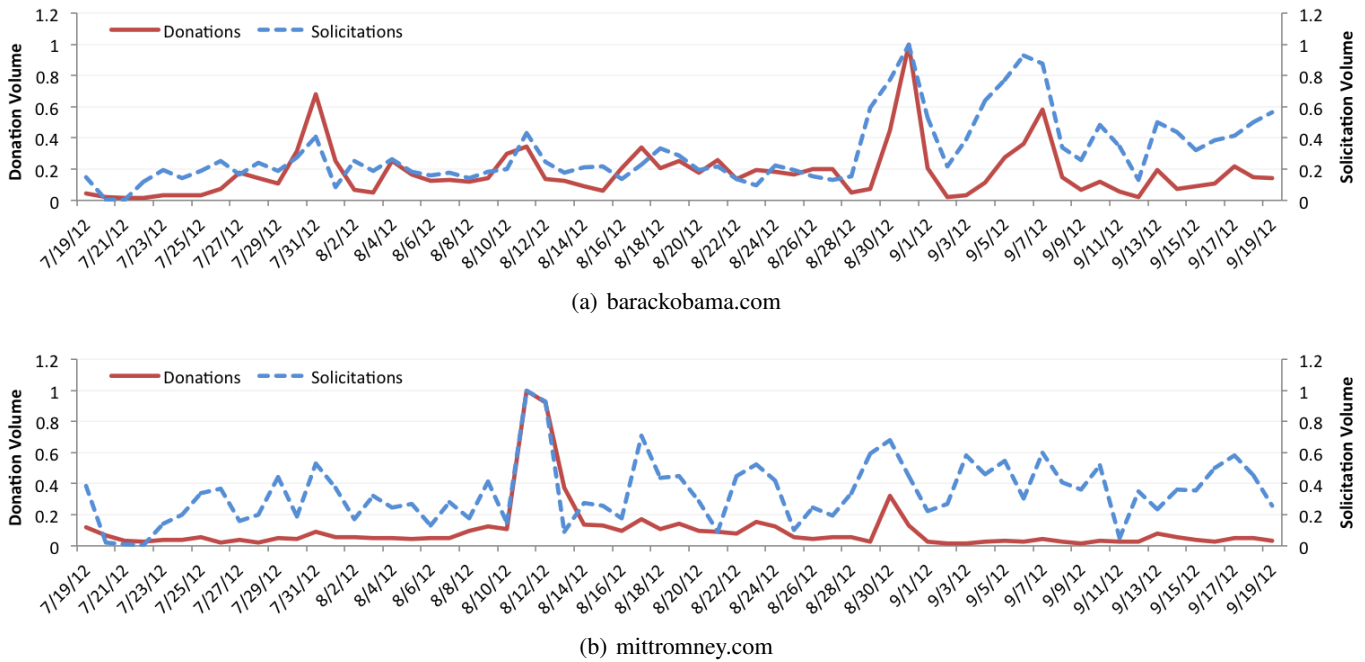
(a) barackobama.com

(b) mittromney.com

Figure 8. Daily donation and solicitation levels for select organizations (1 is maximum volume)

bump. However, the announcement of Paul Ryan's nomination for Vice President on Republican ticket has especially spurred solicitation and donation activity for *mittromney.com*. The national conventions in late August and early September encouraged comparatively few donations on Republican side, whereas spurring donations to *barackobama.com*.

Additionally, we have made the following observations from our manual analysis of select charities:

- Major donations spikes seem to coincide with major events, especially in the political arena.

- Explicit solicitations for donations spur more donation activity than newsletters or updates.

- Solicitations connected to real-world events are more effective than general-purpose ones.

However, a more comprehensive study linking world events, media coverage, and donation behavior is needed to qualify the effect of new information on donors behavior. We leave such a study for future work.

REGRESSION ANALYSIS

Finally, we use a regression model to examine the contributions each of the dimensions we discussed make to the likelihood of a donation. For each user/charity pair we compute the independent variables described in Table 4. These include simple behavior measures (emails sent and received), demographics, interest, and information about the friends of the user. Note that it was impossible to include the external influences such as world events as a factor. To address multi-collinearity issue, we compute variance inflation factors for each variable. These range between 1.01 and 1.72, all within acceptable range for inclusion in analysis. The most related

variables, as shown by Pearson correlation of 0.65, are % bachelors degrees (D3) and median household income (D4). Finally, the dependent variable is whether the user donates to a charity – one if the user has donated to a charity, and zero otherwise.

Because of the prohibitive data size, we devise a sampling strategy that is designed to mitigate the influence of the charities which have more representation (such as those on political and medical issues). Within each charity type we collect 1000 positive examples (those who have donated) and 1000 negative examples (from both interested and general groups). Table 5 shows the logistic regression coefficients, odds-ratios associated with them (exponentiated coefficients), and the mean and standard deviation statistics for each variable. All coefficients, except for I4 and S4 are significant at $p < 0.01$. For that reason it is most informative to consider the odds-ratios listed in the third column.

The effect of S1 (number of friends who are donors to the charity) is by far the greatest: for each friend who donates to a charity, the likelihood for the user to also donate to that charity goes up by 24%. These results strongly attest to homophily – a tendency of individuals to associate with similar others. Simply, you are likely to have similar interests and values as your friends, and thus donate more to the same causes. To gauge how well the model fits the data, we calculate McFadden's pseudo R^2 to be 0.319. Since this measure tends to be smaller than R^2, values of 0.2 to 0.4 are considered highly satisfactory [45].

However, the importance of these factors varies across charity groups. S1 remains to be the most powerful predictor in most cases, being strongest for Religious (87.6%), Humanitarian (49.7%), and Animals (32%). The number of emails received

Table 4. Regression variables

Behavior	
B1	# emails received
B2	# emails sent
Demographic	
D1	age (in years)
D2	gender (male = 1)
D3	% bachelors degrees
D4	median household income (in thousands of dollars)
Interest	
I1	# emails received from charity
I2	# emails received from any charity within a topic
I3	# emails received on topic judging from subject
I4	# emails sent on topic judging from subject
Social	
S1	# friends who are donors to the charity
S2	# friends who are donors to any charity on this topic
S3	# friends who got emails from the charity
S4	# friends who got emails from any charity on this topic

Table 5. Logistic regression of Behavior, Demographic, Interest, and Social factors on Donation (all coefficients are statistically significant at $p < 0.01$ except for I4 and S4)

	Coefficient	Odds-ratios	Mean	SD
intercept	7.83e-02	-	-	-
B1	2.23e-06	1.0000	1079	1351
B2	2.04e-06	1.0000	179.6	725.9
D1	1.95e-04	1.0002	46.35	14.96
D2	-1.04e-02	0.9897	0.419	0.493
D3	3.74e-04	1.0004	29.92	12.87
D4	5.38e-04	1.0005	67.07	24.39
I1	7.70e-03	1.0077	3.847	16.973
I2	-4.83e-04	0.9995	16.24	47.27
I3	5.25e-05	1.0001	15.26	50.64
I4	-3.14e-04	0.9997	0.053	1.822
S1	2.15e-01	1.2403	0.074	0.370
S2	-1.74e-02	0.9827	0.210	0.643
S3	1.48e-04	1.0001	0.841	17.57
S4	1.66e-05	1.0000	9.881	64.04

from charity (I1) is a strong predictor for Children at 8.9%, Military (5.8%), Humanitarian (4.5%), and Animals (4.4%). Gender (D2) is also important, with being male increasing the likelihood of donation by 8.2% to political causes, but decreasing by 5.1% for medical ones. Also, the number of donors to any charity on the topic (S2) is important to Children at 8.0% and Broadcasting at 7.6%. These findings illustrate the difference in behavior of the potential donors to these varying causes, and the need to adjust one's campaign to its audience.

DISCUSSION

Our findings are two-fold. First, the donation behavior we uncovered is consistent with several theories about motivation and giving. Close social connections play an important role in whether an individual donates to a cause, supporting findings of earlier studies [8]. Solicitations, as far as they make individuals aware of issues and charitable opportunities, indeed seem to have some effect on the donation behavior, and considering the low price of sending emails, may be a

great way to promote one's cause. This confirms the popularity of email as a form of communication with organizations by donors found in [10]. Second, we illustrate the power of email communication to reveal human behavior as it records the daily interactions of individuals with the outside world. In this particular domain, the behavioral and social signal is strong enough to effectively predict online donation.

In light of the results above, we offer a set of general recommendations for an effective solicitation campaign. First and foremost, our findings show that such campaigns can be effective, and that users read and respond to the solicitations received via email (confirming general positive reaction to emails found in [10]). Similar to the conclusions in [38], these solicitations are even more effective if the individual is already interested in the topic, as in the case of political or religious issues. Our findings suggest that campaigns should try to make appeals more salient to prospective donors by capitalizing on relevant current events.

Furthermore, demographic targeting may benefit individual campaigns. However, in general, we find that who your friends are is a far better indicator of your propensity to donate than who you are as measured by demographic characteristics. Whether appeals from friends to friends are particularly effective we cannot say, but we have demonstrated a great degree of homophily such that appealing to a donor's friends by any means available (most likely through the donor herself) is likely to be more effective than an appeal to particular demographics (in contrary to the findings in [41]).

Finally, we caution that one size does not fit all. While we make a set of general recommendations, to determine the causal effect of any given email or email campaign, charities must run randomized, controlled experiments in which prospective donors on an email list are randomly assigned to a treatment group (an email or email campaign) or a control group (no emails), thereby allowing them to measure the difference in donation behavior between the two groups and estimate a causal effect of a particular solicitation on opening an email, click-through rates, and ultimately donation (see [16] for more details on running this kind of field experiment).

Although the size and diversity of our dataset provides a unique view of the interactions between donors and charitable organizations, there are several notable limitations to our study. As illustrated by the case of public broadcasting organizations, there are ways other than email in which the information about a cause may travel, such as via radio, television, or other forms of advertisement. Also the link between social ties and donation behavior suggests a communication of this information within the social network, possibly outside email. However, more research should be done on the effect of charity's email on a potential donor. As we mention earlier, the actions of the donor (such as making a donation) may affect the future behavior of the charity (prompting more frequent requests). Finally, our study encompasses donation behavior as witnessed in an email exchange of select users. The interactions between donors and charities outside email were not captured here. Self-selection dictates that the users in our study are more comfortable with internet and email than

those who donate "offline", and may have other peculiarities we have not accounted for. However, the demographic similarity of the donors to the US Presidential campaigns to the election exit poll statistics shows the efficacy of email mining for monitoring social phenomena.

Particular to email collections is the importance of the role the email service plays in users' interactions with the email. As discussed in Spam Filtering section, the spam tagging and the routing of emails into various folders affects the final user experience with the charitable campaign. Furthermore, user-defined filters introduce variables which are difficult to incorporate into data aggregation and analysis. Thus, the policies and spam classifiers of email providers may have an unintended effect on the success of a solicitation campaign. It is obvious, then, that charities are to make sure not to spam their potential donors, as is the case for *wikimedia.org* – the questionable messages were obviously about unrelated content (with subjects like "Wonder pills for thrills" and "Check out this hot babe"), and probably come from a malicious source other than the charity. In fact, the actual donations come from *donatewikimedia.org*, whereas the others from nonsensical accounts, indicating a faked "from" field.

CONCLUSIONS

Our paper provides a unique exploration of the donation behavior of hundreds of thousands of users. Unlike most previous studies on motivation and charitable giving, we take a data-driven approach in which we examine the incoming and outgoing emails for these users over the span of two months. These two months encompass several notable world events, numerous donation drives, and hundreds of organizations. We attempt to quantify the effects of four major factors on donation behavior: demographic, interest, social network, and other external influences. By examining each factor separately and all together we show the contribution of each to the individuals' motivation to donate to various causes.

Most notably, we find a strong relationship between social ties and donations, with the odds of user donating increasing dramatically with every friend who has donated to the same cause.

However, this study provokes more questions for future research. What messages and solicitations are the most convincing to donors? What strategies can charities employ to find new potential donors? And, expanding the data sources, what other technologies support the communication between organizations and the public that leads to further donations? In an increasingly online world, the answers to these questions will determine which charities succeed and which fail at soliciting donations to help their cause.

REFERENCES

1. Adamic, L. A., and Glance, N. The political blogosphere and the 2004 U.S. election: divided they blog. In *LinkKDD* (2005), 36–43.

2. Akerlof, G., and Kranton, R. *Identity economics: how our identities shape our work, wages, and well-being*. Princeton Univ Pr, 2010.

3. Andreoni, J. Giving with impure altruism: Applications to charity and ricardian equivalence. *The Journal of Political Economy* (1989), 1447–1458.

4. Andreoni, J. Impure altruism and donations to public goods: a theory of warm-glow giving. *The Economic Journal 100*, 401 (1990), 464–477.

5. Batson, C., and Shaw, L. Evidence for altruism: Toward a pluralism of prosocial motives. *Psychological Inquiry 2*, 2 (1991), 107–122.

6. Bekkers, R. George gives to geology jane: The name letter effect and incidental similarity cues in fundraising. *International Journal of Nonprofit and Voluntary Sector Marketing 15*, 2 (2010), 172–180.

7. Bekkers, R. Who gives what and when? a scenario study of intentions to give time and money. *Social Science Research 39*, 3 (2010), 369–381.

8. Bekkers, R., and Wiepking, P. A literature review of empirical studies of philanthropy. *Nonprofit and Voluntary Sector Quarterly 40*, 5 (2011), 924–973.

9. Bennett, R. Impulsive donation decisions during online browsing of charity websites. *Journal of Consumer Behaviour 8*, 2-3 (2009), 116–134.

10. Bhagat, V., Loeb, P., and Rovner, M. The next generation of american giving. *White paper* (2010).

11. Chan, M. The impact of email on collective action: a field application of the side model. *New Media & Society 12*, 8 (2010), 1313–1330.

12. Cheung, C., and Chan, C. Social-cognitive factors of donating money to charity, with special attention to an international relief organization. *Evaluation and Program Planning 23*, 2 (2000), 241–253.

13. Collins, K. Who gives? political messages, activist motivations, and campaign contribution behavior. *Working paper* (2012).

14. Conover, M., Ratkiewicz, J., Francisco, M., Gonçalves, B., Menczer, F., and Flammini, A. Political polarization on twitter. In *ICWSM* (2011), 89–96.

15. Eubank, S., Guclu, H., Kumar, V., Marathe, M., Srinivasan, A., Toroczkai, Z., and Wang, N. Modelling disease outbreaks in realistic urban social networks. *Nature 429*, 6988 (2004), 180–184.

16. Gerber, A., and Green, D. *Field experiments: Design, analysis, and interpretation*. WW Norton, 2012.

17. Goecks, J., Voida, A., Voida, S., and Mynatt, E. D. Charitable technologies: Opportunities for collaborative computing in nonprofit fundraising. In *Proceedings of the 2008 ACM conference on Computer supported cooperative work*, ACM (2008), 689–698.

18. Gomes, R., and Knowles, P. Strategic internet and e-commerce applications for local nonprofit organizations. *Journal of Nonprofit & Public Sector Marketing 9*, 1-2 (2001), 215–245.

19. Greenberg, M. D., Hui, J., and Gerber, E. Crowdfunding: a resource exchange perspective. In *CHI* (2013), 883–888.

20. Greenberg, M. D., Pardo, B., Hariharan, K., and Gerber, E. Crowdfunding support tools: predicting success & failure. In *CHI* (2013), 1815–1820.

21. Hart, T. ephilanthropy: Using the internet to build support. *International Journal of Nonprofit and Voluntary Sector Marketing 7*, 4 (2002), 353–360.

22. Hayes, B. M. Everyday information in american philanthropy: informed giving. In *Proceedings of the 2011 iConference*, ACM (2011), 680–681.

23. Hill, S., Provost, F., and Volinsky, C. Network-based marketing: Identifying likely adopters via consumer networks. *Statistical Science 21*, 2 (2006), 256–276.

24. Hsieh, G., Hudson, S. E., and Kraut, R. E. Donate for credibility: how contribution incentives can improve credibility. In *Proceedings of the 2011 annual conference on Human factors in computing systems*, ACM (2011), 3435–3438.

25. Karlan, D., and List, J. Does price matter in charitable giving? evidence from a large-scale natural field experiment. *The American Economic Review* (2007), 1774–1793.

26. Kuo, P.-Y., and Gerber, E. Design principles: crowdfunding as a creativity support tool. In *CHI* (2012), 1601–1606.

27. Liu, Y., Chen, R., Chen, Y., Mei, Q., and Salib, S. "i loan because...": understanding motivations for pro-social lending. In *WSDM* (2012), 503–512.

28. McPherson, M., Smith-Lovin, L., and Cook, J. Birds of a feather: Homophily in social networks. *Annual review of sociology* (2001), 415–444.

29. Muller, M., Geyer, W., Soule, T., Daniels, S., and Cheng, L.-T. Crowdfunding inside the enterprise: employee-initiatives for innovation and collaboration. In *CHI* (2013), 503–512.

30. Mutz, D. *Hearing the other side: Deliberative versus participatory democracy*. Cambridge University Press, 2006.

31. Nuttin, J. Narcissism beyond gestalt and awareness: The nameletter effect. *European Journal of Social Psychology 15*, 3 (1985), 353–361.

32. Olson, M. *The logic of collective action: Public goods and the theory of groups*, vol. 124. Harvard Univ Pr, 1965.

33. Peloza, J., and Steel, P. The price elasticities of charitable contributions: a meta-analysis. *Journal of Public Policy & Marketing* (2005), 260–272.

34. Pepitone, J. Text donations raise $7m for red cross haiti effort. *CNN Money* (2010). `http://money.cnn.com/2010/01/14/technology/haiti_text_donation/`.

35. Rotman, D., Vieweg, S., Yardi, S., Chi, E., Preece, J., Shneiderman, B., Pirolli, P., and Glaisyer, T. From slacktivism to activism: participatory culture in the age of social media. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, ACM (2011), 819–822.

36. Sakamoto, M., Nakajima, T., Liu, Y., and Alexandrova, T. Achieving sustainable society through micro-level crowdfunding. In *CHI* (2013), 1443–1448.

37. Sargeant, A., and Hudson, J. Donor retention: an exploratory study of door-to-door recruits. *International Journal of Nonprofit and Voluntary Sector Marketing 13*, 1 (2008), 89–101.

38. Sargeant, A., and Lee, S. Donor trust and relationship commitment in the uk charity sector: The impact on behavior. *Nonprofit and Voluntary Sector Quarterly 33*, 2 (2004), 185–202.

39. Shaw, K. Defining the self: Constructions of identity in community college students. *Community colleges as cultural texts* (1999), 153–172.

40. Shields, T., and Goidel, R. Who contributes? checkbook participation, class biases, and the impact of legal reforms, 1952-1994. *American Politics Research 28*, 2 (2000), 216–233.

41. Shier, M., and Handy, F. Understanding online donor behavior: the role of donor characteristics, perceptions of the internet, website and program, and influence from social networks. *International Journal of Nonprofit and Voluntary Sector Marketing* (2012).

42. Simon, A. Television news and international earthquake relief. *Journal of Communication 47*, 3 (1997), 82–93.

43. Tajfel, H., and Turner, J. An integrative theory of intergroup conflict. *The social psychology of intergroup relations 33* (1979), 47.

44. Unger, L. Altruism as a motivation to volunteer. *Journal of Economic Psychology 12*, 1 (1991), 71–100.

45. Veall, M. R., and Zimmermann, K. F. Pseudo-r2 measures for some common limited dependent variable models. *Journal of Economic Surveys 10*, 3 (1996), 241–259.

46. Wagner, C., and Wheeler, L. Model, need, and cost effects in helping behavior. *Journal of Personality and Social Psychology 12*, 2 (1969), 111.

47. Wash, R., and Lampe, C. The power of the ask in social media. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work*, ACM (2012), 1187–1190.

48. Weerts, D., and Ronca, J. Profiles of supportive alumni: Donors, volunteers, and those who "do it all". *International Journal of Educational Advancement 7*, 1 (2007), 20–34.