Giving by the Numbers:
A Study of Peer Effects with Online Crowdfunding

A Thesis
Presented to
The Established Interdisciplinary Committee for Mathematics and Economics
Reed College

In Partial Fulfillment
of the Requirements for the Degree
Bachelor of Arts

Claire Jellison

May 2020

Approved for the Committee
(Mathematics and Economics)

Tristan Nighswander

Jerry Shurman

Acknowledgements

I'm so grateful to have gotten to know, work with, and learn from so many incredible people during these past four years.

To my advisors, thank you for the time you dedicated to improving this thesis. Tristan, thank you for always showing so much enthusiasm for this project. Bouncing around ideas in your office was definitely a highlight of this process and I always left our meetings with a renewed sense of determination. Jerry, thank you for making me excited to learn math again.

To the many teachers I've had, I'm grateful. Jonathan, Adam, Andrew, Kyle, Kim, Jeff, Noelwah, Alfonso, and Misha, your classes and dedication have been immensely inspiring to me.

To my housemates, Purna, Sophia, Shawn, and Thomas, there are no other people that I would have rather come home to. I feel so lucky to live in a house where every endeavor is supported, everything is endlessly funny, and nothing is ever clean.

To my econ friends, Zeki, Will, Shulav, Jacob, Ryan, Salma, and many others, I'm so glad to have gotten to know you all. Sitting around and talking at commons for hours on end was a waste of time I'm inevitably going to miss.

To my miscellaneous friends, Sara, Nick, Yuki, Xitlalitl, Patricia, Kayla, Kees, and David, thank you for the commiseration, ping pong games, fireside debates, questionable cuisine, climbing trips, hazardous street crossings, workplace competitions, and all else. You made my time here so much more enjoyable.

Finally, thank you to my parents, who first taught me how to love and question the world.

Contents

List of Tables

List of Figures

Abstract

This thesis investigates the size and nature of peer effects using data collected from the online crowdfunding platform GoFundMe. On a campaign page, the previous five or ten donations are visible to potential donors. This allows for investigation into how the knowledge of these previous donations affects future donors' decision to give. I find that donors appear influenced by the donations that are displayed to them but the size of the effect is significantly different for men, women, and anonymous donors.

Introduction

In microeconomics the decision of how much to spend is relatively straightforward. A person chooses how much to consume of each good by maximizing their utility, or overall satisfaction, subject to their budget constraint. Their utility, in turn, stems from a set of underlying preferences. These preferences are well-defined, transitive, and constant with respect to time. Holding preferences constant and allowing relative prices to vary is how decision making is generally modeled. However, we know that people are not endowed with a set of immutable exogenous preferences and decisions are rarely made in such vacuum-like conditions.

Oftentimes consumption is visible which allows for social elements to enter into our decision calculations. With highly visible goods such as clothing, it's clear that people do not act independently; rather, we often mimic one another, observing others and mirroring their decisions. We are constantly receiving cues about how people are spending their money, and it can influence how we choose to spend our own money.

One of the most well known sociological phenomena is conspicuous consumption, originally introduced by Veblen (1899). According to this theory, people spend disproportionately more money on certain high-quality visible goods, like Rolex watches, because owning these goods allows one to convey status (Bagwell & Bernheim, 1996). Buying certain goods can help us try to craft an identity or attempt to shape the way we would like to be perceived. These social aspects ultimately affect how we make decisions and they are highly malleable, contingent on our knowledge of others' actions or perceptions. Yet this notion is fundamentally at odds with the standard preferences of neoclassical consumer theory; people do not act independently from one another but instead exhibit concern for how their consumption compares to that of their peers. However, as increased access to personal technology allows us to make our lives more visible than ever, conspicuous consumption is unlikely to stay limited to the same traditional goods.

This thesis examines how the social factors that accompany visible consumption enter into the realm of online charitable giving. When charitable donations are publicly displayed, donors may act in a way consistent with conspicuous consumption; they may exhibit concern for how their donation compares to those of their peers. Indeed, charitable giving may even have some advantages over other forms of visible consumption because it may be more socially sanctioned and visible to a larger audience (Glazer & Konrad, 2008).

Through analyzing data from the leader in online crowdfunding, GoFundMe (GFM), we attempt to understand how the visibility of donations impacts the way that people behave. On the website the previous five donations are publicly displayed. This layout creates a natural field experiment where a potential donor is presented with a clear set of information points regarding the donations of others prior to making their own donation. This generates a two-fold effect, people are prompted with the responses of others and they have the knowledge that their own donation will be visible unless they choose to make it anonymously. If future donations appear highly responsive to the displayed donations, this would indicate greater concern for how one's donation compares to those made by others, whereas the absence of an effect of displayed donations would be evidence that people act more independently. Furthermore, if anonymous donations adhere to the same pattern then we might presume that the effect not entirely attributable to prestige related reasons. However, there is not a strong consensus in the economics literature about how and to what extent an individual's donation decision is impacted by the knowledge of other's donations. Therefore, this thesis attempts to not only confirm the existence peer effects, but also to understand in a more concrete sense how people's donation sizes are shaped by the amounts they know others have donated.

In the first chapter, I begin by describing how the social effects tied to visible consumption have been incorporated into contemporary economics. I then turn my attention specifically to models of charitable giving. In particular, I focus on two distinct models that carry divergent implications for how people might respond to the donations of others. The first adopts the more standard neoclassical framework of consumer behavior while the latter allows for donors to experience "prestige" or a "warm glow" that stems from the act of donating. I also outline how the effects described in the second model may be consistent with the interpersonal comparison that accompanies visible consumption. Then, in the second chapter, I detail information about my choice of platform, my method of data collection, and general characteristics of the data. From there, in the third chapter, I proceed to construct a variety of models to test whether donors appear to be responsive to the donations displayed to them. Lastly, I explain how these results fit into the theoretical landscape surrounding charitable giving, conspicuous consumption, and interpersonal comparison.

Chapter 1

Literature Review

There has long been recognition that visible consumption is differs substantially from ordinary consumption. One of the earliest examples is Veblen's *The Theory of the Leisure Class* which outlined how psychological and social components serve as key determinants in the formulation of preferences. According to Veblen's theory, certain visible luxury goods are consumed mainly for the purposes of impressing others. This implies that the utility from these good emerges in part from how these goods compare to the goods consumed by one's peers (Friedman & Ostrov, 2008). While Veblen's ideas took root in sociology, they initially failed to generate substantial interest within economics. Major theories of consumer demand that emerged during the middle of the 20th century such as Friedman's permanent income hypothesis or Modigliani and Brumberg's life cycle hypothesis did little to incoporate social effects, operating under the assumption that such effects would be trivial at the aggregate level (Mason, 2000).

A notable exception, however, was the theory proposed by Duesenberry in 1949 with the publication of *Income, Saving and the Theory of Consumer Behaviour* which bore significant resemblance to Veblen's thinking (McCormick, 1983). Duesenberry argued that existing theories of consumption failed to account for how the level of ones expenditures could be determined not only by changes in income and prices but also by witnessing the consumption patterns of others. This phenomena was driven by the interdependence of people's preference systems whereby consumers would emulate each other in order to preserve or increase their social status and prestige (Mason, 2000). Despite the general neglect of Duesenberry's ideas during the time of his publication, models where consumers care about their relative consumption have recently gained significant traction within several areas of economic thought (Rablen, 2008).

Within models of social comparison, which originated with Duesenberry, outcomes are judged according to compound reference points, such as the societal averages. While with rank based utility, people are concerned with the ordinal rank of their consumption within the broader population (Friedman & Ostrov, 2008). Friedman & Ostrov (2008) proposes a model of interpersonal comparison where consumers exhibit differing motivations. When "pride" is the primary motivator people are concerned with consuming in excess of their peers. Alternatively with "envy", people are upset when their own consumption is short of what their peers are consuming. Such models of interpersonal comparison often distinguish between two categories of goods, nonpositional and positional goods. With positional goods people are concerned with their relative level of consumption whereas with less visible nonpositional goods relative concerns are largely unimportant. The presence of these social effects with positional goods is demonstrated to have a distortionary effect in how people allot their money without cooperation (Frank, 1985).

Given the public nature of donations on GFM and the ability to see past donations, it seems plausible that donations may be a positional good where people are concerned with the relative size of their donation. Furthermore, since the website prompts donors by displaying a set number of previous donations, donors have some awareness of how a potential donation amount would be situated relative to other donations. Thus, the amount they decide to contribute can also give us insight into the motivating forces behind interpersonal comparison. Do donors attempt to outcompete one another or avoid stigma? The potential for observable peer effects in this scenario means that this project falls at the intersection of two lines of inquiry, charitable giving and interpersonal comparison. On the one hand, if interpersonal comparison is an important aspect of charitable donations on the website, we would expect that donors would somehow adapt their donation amount based on the displayed donations. However, charitable donations are not a typical good. Even without the presence of peer effects, we might expect that future donations would depend to some degree on previous donations. Therefore, we now examine the standard models for charitable giving and the lengths to which social effects have been incorporated.

1.1 Models of Giving

A majority of US citizens give to charitable causes each year, with charitable donations now constituting more than 2% of GDP. Moreover, since 1968, the growth in charitable donations has greatly outpaced the growth of the S&P500. Individual donors account for over three quarters of the total number of gifts each year (List, 2011). This means that charitable donations are a non-negligible and increasingly substantial piece of the economy. Yet, beyond the magnitude of the charitable donation sector, the widespread nature of giving makes it worthy of attention especially within economics, a field dedicated to understanding the allocation of scarce resources. The challenge lies in reconciling voluntary wealth transfers with the self-interested utility-maximizing agents that govern microeconomics. The basic question is what makes people compelled to give?

Models used to answer this question generally fall into two categories, one where the benefit from donating is primarily public and the other where the benefit is private. In the case of a public benefit both the donor and other individuals may benefit. For example, when donating to public radio both the donor and others get to listen to the radio station as a result. In contrast, a private benefit is exclusively enjoyed by the donor. This encompasses all the good feelings or abatement of negative feelings the donor might experience from giving. Unlike with the public benefit, with the private benefit social effects may play an important role.

The distinction in the nature of the benefit proves important as it generates divergent implications for how people respond to changes in income or in the price of giving. Vesterlund (2006) notes that if we consider the benefit to be primarily private, then individuals will donate according to how much they value the product of a nonprofit or charity. Conversely, if the benefit is primarily public then someone else's donation provides the same benefits as one's own donation. Given that one can receive the same utility without incurring any cost, people are strongly incentivized to free-ride off of the donations of others. Consequently, when the donations of others increase in size, we would expect an individual's donations to decrease in response. Alternatively, if the benefit obtained by donating is primarily private then the effect of changes in others' donations should not exert as large of an effect on an individual's donation. Nonetheless, we also see that many of private benefits may have a social component thus confounding the relationship between previous and future donations.

1.1.1 Classical Model of Giving

The classical model focuses on the public benefit and assumes that the individuals derive utility from the nonprofit's output and their private consumption of the output, treating the contributions of others as given. In this model, each individual has little incentive to give and would be inclined to free-ride off of others' spending. This, in turn, leads to the possibility of complete crowding out (Vesterlund, 2006).

In this situation, when some agent (though the term is most frequently used with respect to the government) increases spending or funding to a certain service it could trigger a reduction of equal magnitude of private sector funding to support the same service. This is because the desired level of output can now be achieved without the contribution of the individual. Both of these implications, complete crowding out and extensive free riding, seem overly pessimistic and offer a poor depiction of what occurs in practice. Overall, the classical model bears little resemblance to what is observed in the empirical setting with charitable giving (Glazer & Konrad, 2008). Alternative models, which incorporate various private benefits, have been proposed and include avenues to reduce the propensity to free-ride without leading to the extreme outcome of complete crowding out.

1.1.2 Private Models of Giving

Private models of giving allow donors to receive a personal benefit that stems from the act of donating. That is, part of the benefit derived from charitable giving comes from making the donation rather than from the good that it buys. The private model no longer leads us to anticipate dollar-for-dollar crowding out when others donate to a cause that we are interested in; there is an added benefit to making the donation yourself that can't be experienced from another's donation. Further, these models provide a framework in which the incentive to free-ride is substantially weakened for the same reason.

The nature of the personal benefit that a donor obtains has received considerable attention within the charitable donation literature resulting in a large swath of theories designed to account for this donor-specific benefit. Harbaugh suggests that giving may enter the donor's utility function through two distinct mechanisms, the "intrinsic benefit", the donor's own knowledge of what they have given, and the "prestige benefit", which comes from other people witnessing the donor's generosity (Harbaugh, 1998). However, Harbaugh notes that the prestige benefit and the intrinsic benefit may not enter directly into the donor's utility function but rather as an amount relative to the gifts made by others. This would mean that the effect's magnitude would be dependent on the donor's knowledge of the actions of others rather than absolute. In an empirical study of lawyers' donations to their law school, Harbaugh finds evidence that the prestige benefits accounts for a significant portion of donations. Glazer & Konrad (2008) construct a model where donors contribute in order to signal their wealth as with the theory of conspicuous consumption. Furthermore, they note that charitable donations may have some advantages over private goods, because unlike some forms of conspicuous consumption they may be less prohibited by social norms and can be visible to peers that a person does not directly come into contact with.

At the more extreme end of the private models of giving, Tullock suggests that the donor is not directly interested in the well-being of others. Rather, charitable giving allows one to purchase a feeling of satisfaction derived from one's financial sacrifice. This satisfaction in part stems from the benefit to the donor's reputation. He notes that the benefit to reputation is rooted in others' perceptions of the charity's image rather than its true effectiveness. This means that the gift enhances one's reputation when the people who matter to the donor consider the cause to be a worthy one (Tullock, 1966).

The "intrinsic benefit" that Harbaugh comments on is commonly referred to as the "warm glow" in the literature. Models including the so-called "warm glow" have proven to be more consistent with empirical observations and can also help explain diminished free riding and incomplete crowding out (Andreoni, 1989). Furthermore, Romano & Yildirim (2001) demonstrate that if donors only care about the provision of a public good, then a charity has no incentive to announce donations sequentially. However, by doing so, they convert the simultaneous donation game into a sequential game. For agents with standard utility functions, seeing that more people had already contributed would likely reduce or discourage their contribution, however, for donors who experience a private benefit the announcement could encourage them to give.

Through allowing an individual to derive a benefit that is not associated with the output of the charity, these models ease the extreme implications of the classical model. In particular, we see that within the private models of giving, since donors receive an added benefit from the act of donating we would not expect their donations to decrease as the donations of others increase to the same extent as in the classical model. Moreover, within private models, donors are sometimes concerned with how their donation compares to that of their peers (Harbaugh, 1998) or how their donation may impact their reputation (Tullock, 1966). This implies that the size of the donor specific benefit may be dependent to some extent on interpersonal comparison, meaning that knowledge of others' donations would exert influence on how much a donor decides to contribute.

These models provide us with a useful framework for understanding how future donations might respond to previous donations. If subsequent donations decrease in response to previous donations, or the presence of more seed money, this would support the classical model. Alternatively, if subsequent donations are relatively unaffected by previous donations or displayed donations we might assume that there is an additional private benefit from donating that is independent of relative concerns. Finally, if donations actually increase with displayed donations then this may support the idea that the donor-specific benefit may enter the donor's utility function as an amount relative to other known donations.

1.2 Similar Work

The most closely related paper to my thesis estimates the size of peer effects using a data set from two websites that people used to raise money for charity by running the 2010 London marathon (Smith, Windmeijer, & Wright, 2014). This paper looks at how very large donations, very small donations, the mode of past donations, and the mean of past donations affects subsequent donations. They find strong evidence for the presence of peer effects utilizing indicator variables for high and low donations as well as a linear-in-means model. Ultimately, they believe that donors give what they "think that they personally are expected to give where the distribution of the donations of their peers (along with other factors, such as income and specific cause) feed into the formation of that expectation". The difficulty is that because of the nature of their data, the mean of past donations is not really a consistent entity. As more donations arrive, each donation has less leverage in pulling the mean in either direction. They cite this as a finding — the later donations occur, the less influential they are. However, there is an implicit assumption through using the cumulative mean that the sample size does not matter; the mean is treated the same by people regardless of how many donations there are even as it becomes more difficult to compute. However, with their platforms, there is no clear indication of which donations people actually see before donating.

This GFM dataset is well suited for studying peer effects and social comparison because a constant number of donations are always displayed and the user must make an extra effort to see more donations. Furthermore, there are many facets of the site that allow for interesting comparisons. For one, the website initially displayed ten donations and later switched to only display five donations, which allows one to see the effect of an increased number of reference points. Secondly, one can make inferences about anonymity, gender by coding first names, and even potentially familiarity between donors since campaigns range from small scale local efforts to national movements. Thirdly, the immense size of the dataset makes results more reliable and allows for more robustness checks.

A recently published paper (Sisco & Weber, 2019) appears to be the first to analyze a portion of the massive amount of data available on GFM. This study includes internet data from 9,264 campaigns to test several psychological hypotheses regarding the origins of altruism, specifically sexual selection and kin selection. According to the theory of sexual selection, certain traits may be selected for not because they increase one's likelihood of survival but rather because they serve in attracting mates thereby increasing reproductive prospects (Darwin, 1871). Alternatively, with kin selection, traits may be selected for that do not directly increase an individual's likelihood of survival, but that increase another's chance of survival who is likely to have the same genes (Stewart-Williams, 2015). Leveraging donor names to extract gender, they find evidence that if the proportion of visible females on screen increases, female donors give less while male donors give more. They use this to support the theory of sexual selection. Alongside gender, the authors also attempt to measure effects stemming from familial ties. They find that the average donation amount is significantly greater when a donor shares the same last name with the donation receiver and use this to support the theory of kin selection. In relation to peer effects, they find that the mean visible donation on screen at the time of giving was a significant and positive indicator of the value of donations of both men and women, with the effect being stronger in men. Their data analysis provides tentative support for the presence of social comparison in our setting, but this is not the focus of their study. Therefore, this project builds off of the work of Sisco & Weber (2019) in some more minor ways (analysis of anonymity, seed money, controlling for seasonality, time, and campaign specific features etc), but most notably attempts to contribute to understanding *how* people respond to known donations.

Chapter 2

Data Collection

2.1 GoFundMe

My main criteria when searching for potential crowdfunding platforms was evidence that they provided a clear anchor people would see before making a donation, there was a timestamp on each donation, and the site allowed scraping. The companies that I considered include GoFundMe, Crowdrise, Kickstarter, Indiegogo, Patreon, Fundly, and Kazoo. Kickstarter, Indiegogo, and Patreon were quickly ruled out because they focus strongly on a tiered donating approach, with several specified levels at which someone may choose to donate. Additionally, the donor may receive products or certain benefits based on their donation level. Therefore, they are out of scope for a project primarily focused on anchoring within the context of charitable giving. On the other hand, GoFundMe, Crowdrise, Fundly, and Kazoo all appeared to be good potential candidates. I chose GoFundMe because it's the largest of all the sites considered and allows spiders to scrape the majority of its website as indicated in the website's robots.txt[1]. GoFundMe displays the previous five donations[2] to potential donors and the donation history data for a given campaign is retrievable.

GoFundMe claims to be the largest crowdfunding platform. Since its conception in 2010, the site had amassed over five billion dollars from 50 million people through 2017, when fundraising totals were last announced (Monroe, 2019). It's made headlines on numerous occasions, frequently appearing in the wake of national tragedies or alongside political movements. Initially, GFM charged a five percent commission, but

[1] Visible at https://www.gofundme.com/robots.txt, which specifies to web crawlers what actions they are allowed to preform on the website.

[2] While the site now shows the previous five donations, the site used to display the previous ten donations. There is further discussion of this in the data description section.

since 2017 has waived that fee for fundraisers in the US, although there is still a payment processing fee (Heller, 2019). There are 18 categories on GoFundMe including medical, memorial, emergency, non-profit, education, animals, business, community, competition, creative, event, faith, family, newlywed, sports, travel, volunteer, and wishes. The category is designated by the campaign organizer. Medical fundraisers are the most common, amounting to about a third of all campaigns (Monroe, 2019). The layout of the a typical campaign page can be found in the second appendix.

2.2 Web Scraper

I built the GoFundMe web scraper in python using the scrapy package, an open source web crawling framework designed to move through web pages by following specified links in the html source code and extracting data. Given that GoFundMe designates the order in which to display campaigns by some unknown algorithm, one of the initial questions was how to select data from the website at random. Sampling at random from all possible campaigns would ensure that the data was not biased by the characteristics that the website's algorithm selects for and promotes.

In order to counter the potential bias of the website's internal algorithm for ordering campaigns, I initially built one scraper that crawls through the pages of campaigns in a certain category collecting all the displayed links to each campaign page. However, after 1,000 unique campaigns in a given category, no more would be displayed by the site. Given the size of the website, this was clearly a very small subset of all campaigns. My impression is that campaigns that are selected to appear under the category labels, in general, have lots of donation momentum, high goal amounts, and polished profiles. In contrast, through searching certain keywords rather than just following category links, the disparity is evident; many campaigns are far from funding goals and have received little to no attention. Therefore, sampling from the category pages would be highly subject to the logic underlying the site.

Using the search feature provided a way to bypass the website's campaign ordering logic. The search feature on the website allows a viewer to input certain keywords and then pulls up relevant campaigns that use that word. The issue once again, was that while a certain term might say that it yielded millions of results, only around 1,000 would actually be displayed out of the full set of results.

The key to resolving this issue was by using zip codes. Within the search feature, it is possible to specify the country as the US and also include a specific zip code. This allowed me to build a scraper that searches a given zip code on the site and then scrapes all the campaign urls for that zip code. This works because every campaign on the site is necessarily tied to a zip code and thus the full list of zip codes partitions all US campaigns.

I downloaded a full list of US zip codes[3]. I randomly partitioned all the zip codes. I passed this list of zip codes from the first several partitions to the scraper, which traverses the website according to the breadth first search algorithm using three distinct parsing methods. The homepage of the site serves as the root for our graph search. From there, the scraper first completes the search for each zip code from the homepage. These links for each zip code are sent to the first parsing method which pulls all the urls for campaigns attached to that zip code. These links go to the second parsing method which records several campaign specific traits and generate links to the donation history data. These links are sent to the third parsing method which calls itself recursively until all past donations have been recorded. An outline of the scraper can be seen in Figure 2.1 [4].

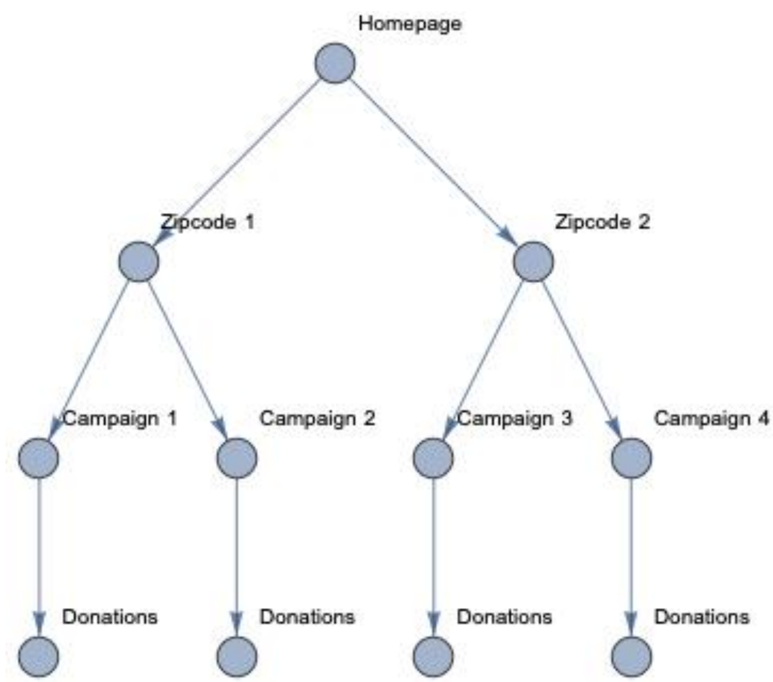

Figure 2.1: Schema of scraper

[3]https://www.unitedstateszipcodes.org/
[4]Source code for the scraper can be found on my github account, https://github.com/zumiko/thesis

2.3 Data Description

The campaign zip code table provides the zip code for every campaign scraped and a unique key for identifying each campaign. The time history data for each campaign includes the campaign key, the amount donated, the time of the donation, the donation identification number, whether the donation was anonymous[5], the self entered name of the donor, as well as the fields “isoffline”, “profileurl”, and “verified”. The campaign attributes data table includes the key for identifying the campaign, the title of the campaign, the category of the campaign, the date that the campaign was created, whether the campaign has been terminated, the campaign organizer, the campaign description if provided, whether the campaign is a registered tax deductible non-profit, the amount raised, and the goal amount.

A total of 15,672 campaigns were collected, with a total of 642,644 individual donations. No more than 1,000 donations could be scraped from a single campaign. The mean donation amount was around $112 and the median was $50. For anonymous donors, the mean donation amount is significantly higher at $127. Approximately 20 percent of donations were made anonymously. Of all campaigns, around 7.3 percent were tax deductible nonprofits. A quick summary of the data is given below in Table 2.1 and the distribution of amounts donated can be seen in Figure 2.2. The mean donation amounts across categories can be seen in Figure 2.3. Furthermore, we see in Figure 2.4 that the first few donations appear to be much higher than subsequent donations. This is of little surprise given that we would expect that the first few donors might have a special relation to the recipient that would cause them to contribute higher amounts. In response to this we will exclude the first five donors from all our models.

	All donations	Female	Male	Anonymous	Unknown
mean donation	112.36	75.83	110.98	126.78	307.73
total donations	642,644	286,658	189,416	120,538	46,032

Table 2.1: Summary of data

In order to gain insights into how gender interacts with charitable giving, first names were assigned a gender, male or female, based on their likelihood of being that gender according to historical data using the gender package in R (Mullen, 2019). Some observations could not be classified and are listed as unknown. These are due

[5] While the name of an anonymous donation is not publicly displayed it is visible to the campaign organizer.

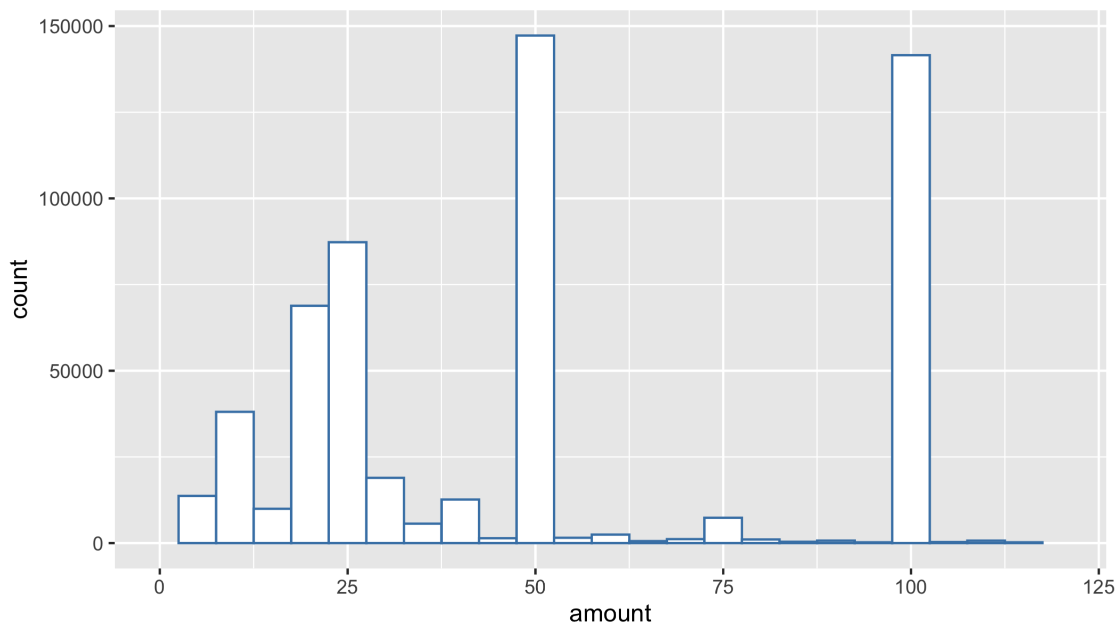

Figure 2.2: Distribution of donation amounts

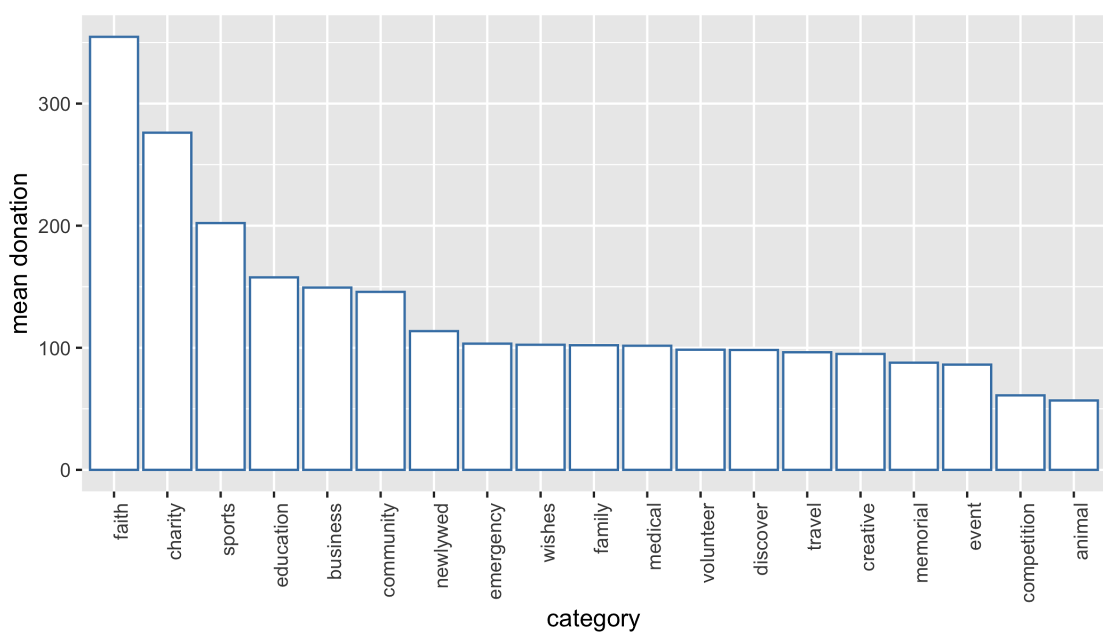

Figure 2.3: Mean amount given by category

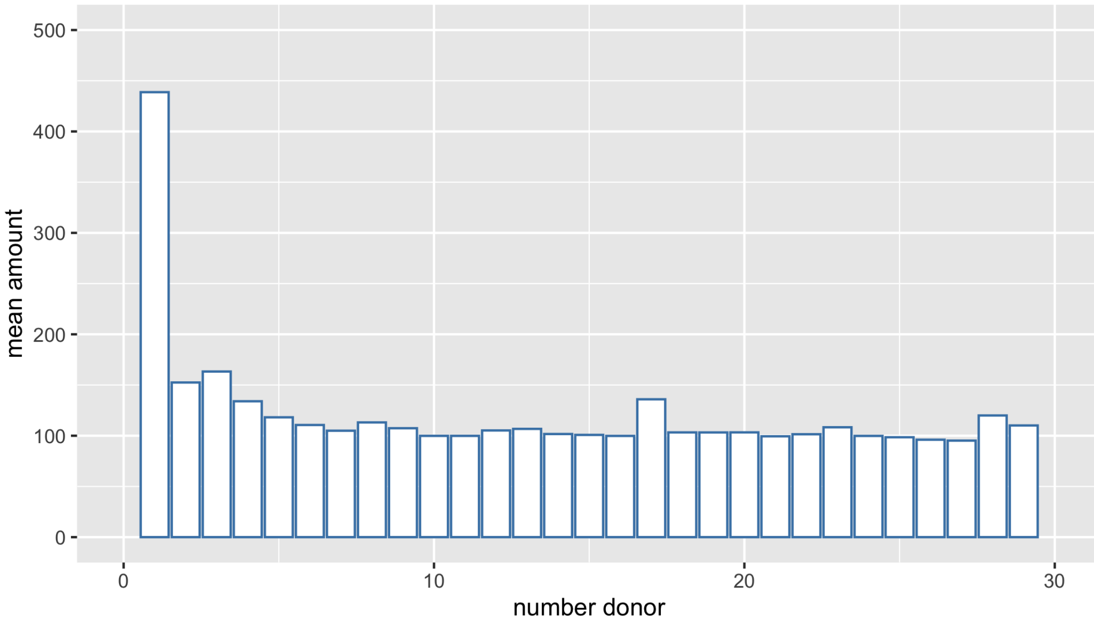

Figure 2.4: Mean amount given by donor number

to names with ambiguous gender, anonymous donations, couples donating (i.e. "Mr. and Mrs. Smith"), donors using their relationship to the recipient (i.e. "Grandma"), or other variants. Obviously, this method is not ideal and thus any interpretation of the results relating to gender should be regarded with mild suspicion.

It is likely that GFM usage may not be uniform across the US and therefore the data may be more representative of certain areas of the country based on where the website is more popular. Furthermore, the demographics of who uses the site is not well known. However, the website is used quite broadly across the US as we can see in Figure 2.5 which plots the location of where campaigns are located based on their zip code. While campaigns have zip codes associated with them, the origin of where a donation was made from cannot be determined from the data.

Another drawback of the data is that the exact date that the website switched from displaying the previous ten lags to the current layout where the previous five lags are displayed is not precisely known. Identifying when the layout of the website changed posed some difficulty. Given the previous study Sisco and Weber (2019) who collected data in June 2016, the layout change occurred sometime after that month. Therefore, all donations prior to that month were coded as having 10 lags. Furthermore, I was able to find a screenshot of the website on July 25th, 2019 with 5 displayed donations, so all donations after that date are coded as having 5 lags[6].

[6] I aso recently found a screenshot for a campaign with ten lags on April 1st, 2019. Using this

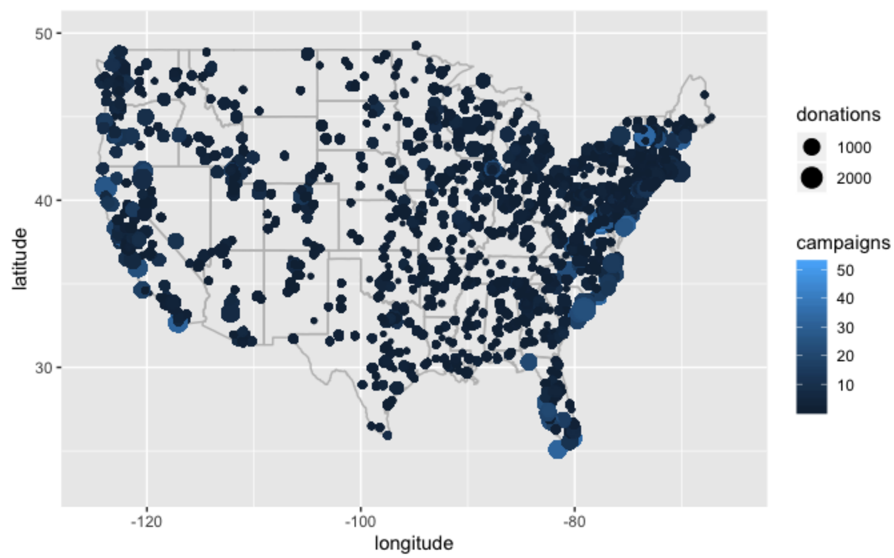

Figure 2.5: Map of data on contiguous states

Therefore, a large portion of the data had to be disregarded for lag analysis given that the displayed number of lags was not known.

One might also wonder how often campaign goals are met. For all campaigns, a percent of goal met was calculated by dividing the amount raised by the goal amount. This distribution of campaigns' progress towards their goal amounts can be seen in Figure 2.6. This is particularly relevant to the work on seed money by List & Lucking-Reiley (2002). Using three levels of seed money (10%, 33%, and 67%), they find at higher levels of seed money both the total contributions to a cause and the size of the donations increase. They identify two potentially conflicting effects as seed money increases. On one hand, donors may be more compelled to free ride off of the donations of others since the fundraiser is more likely to reach the goal even without their contribution. This effect is aligned with the idea that the benefit is primarily public; a donor is free to enjoy the output of the charity reaching its goal regardless of whether or not they played a part in attaining it. Alternatively, there may be a "follow-the-leader" component, where additional seed money sends potential donors a signal of quality of the fundraiser as in the leadership model of giving (Andreoni, 2006). From the discrete amounts of seed money they use in their study, it appears that the "follow-the-leader" effect entirely subsumes the propensity to free ride. However, with a wider range of seed amounts, the interaction between the two effects is not well-known.

With the nearly continuous amounts of seed money on GFM, we can help to

date would substantially increase the sample size of this project.

resolve this question. Since the amount of seed money is necessarily greater for each subsequent donation, we would expect fewer campaigns to stagnate at levels closer to their goal amounts. That is, if the probability of donations increases with additional seed money, we would expect that the number of campaigns to stagnate at each level of percent raised to decrease, which appears to be the case in the figure. This implies that as a campaign approaches its goal amount it becomes more likely to receive more donations, donations of greater size, or some combination of the two. Additionally, we also observe a peak of campaigns at the point where their goal has been fully met demonstrating that the goal amount does have a meaningful effect, perhaps prompting donations as a campaign approaches the goal amount or dissuading donations once the campaign has surpassed it. This will be further examined in the next chapter.

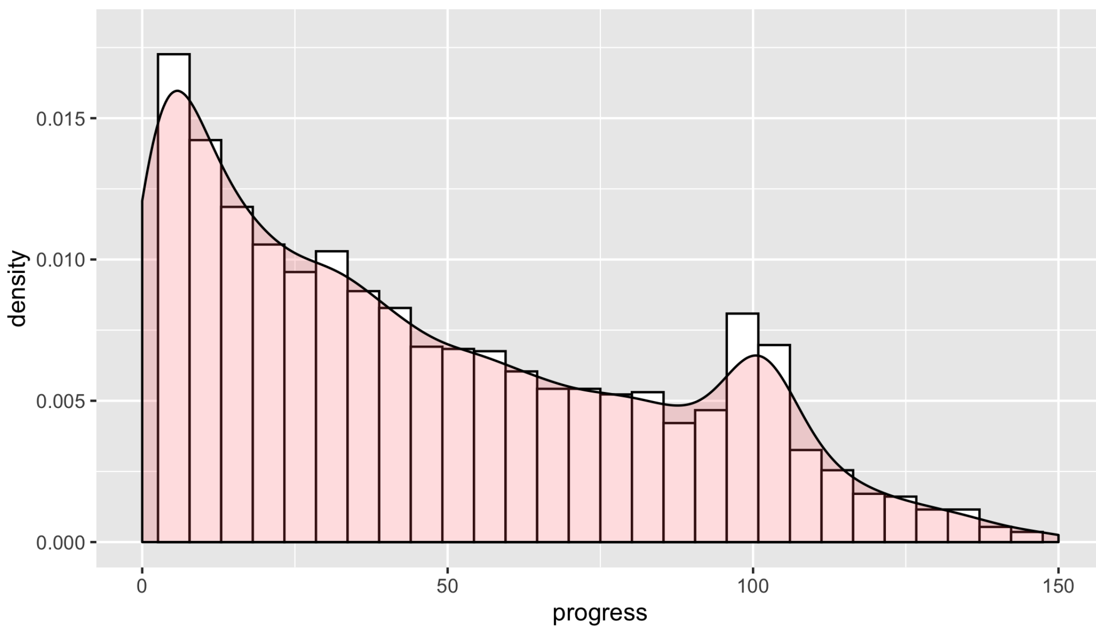

Figure 2.6: Distribution of percent funded

Another thing that should be considered is whether donation amounts are strongly correlated with time of day, day of week or month. If, for example, donations amounts change significantly with the time of day, that could cause nearby donations to appear correlated. Therefore, from the timestamp, the time of day, day of week, month, and year were all extracted. In Figure 2.7, we see that there does not appear to be an obvious trend across the months of the year, although there is a slight decline in December perhaps due to discretionary incomes being used for holiday spending. We also witness that donations appear to have a clear positive skew, which is also evident in the fact that the mean donation is double the median donation. There also does not appear to be strong pattern was for a similar plot with the days of the week on the horizontal axis. Furthermore, in Figure 2.8, we see a scatterplot of donations by their amount and the hour of the day when they were made. Given the volume of data, points that are located very close together are grouped and colored according to the density of nearby points with lighter points displaying a greater density of donations in that location. Here, there is no discernible evidence that higher donations tend to arrive at a different part of the day than lower donations. Rather it appears that donations, regardless of size, arrive when people are awake.

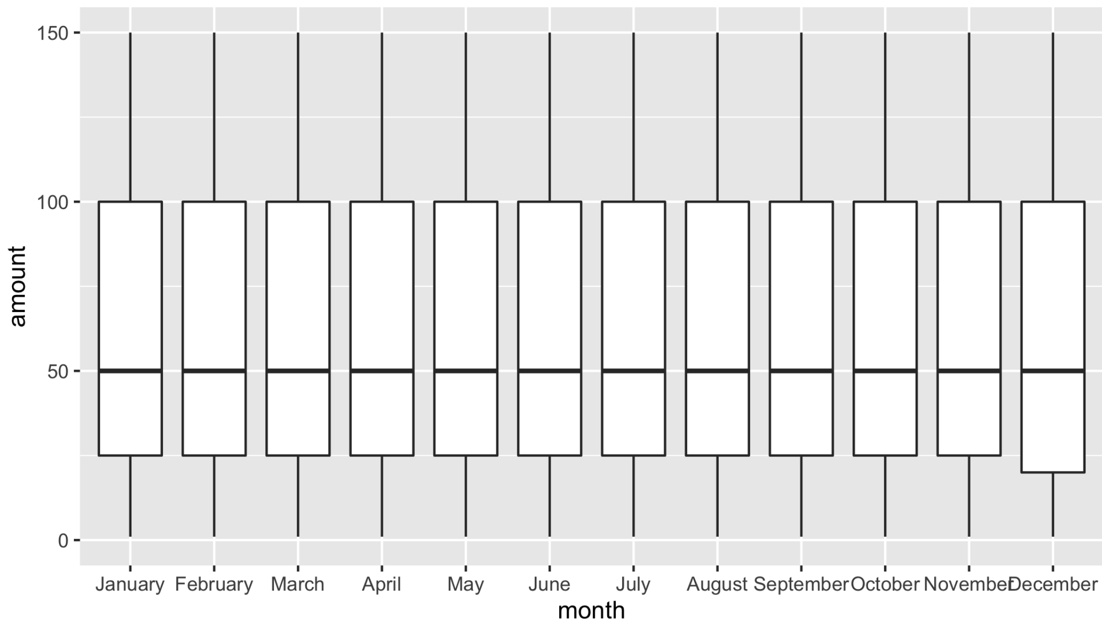

Figure 2.7: Distributions of donation amounts by month

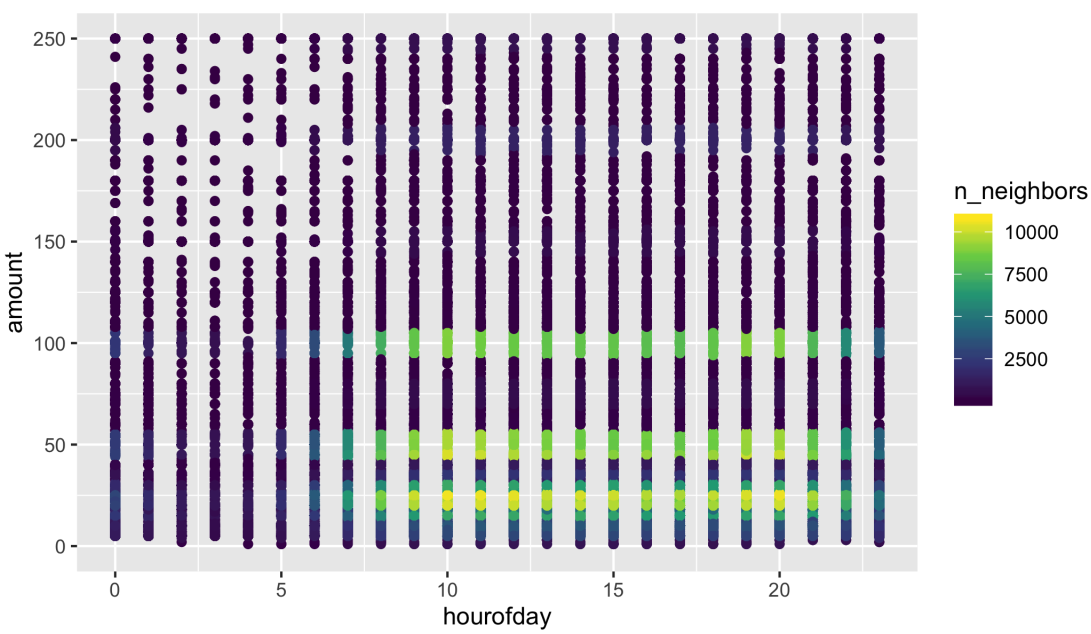

Figure 2.8: Distributions of donation amounts by hour of day

Chapter 3

Regression Models

There is substantial economic literature on how social context can affect individuals, yet empirical studies to measure the strength of peer effects have proven to be of some difficulty (Moffitt, 2000). In general, researchers build regression models with group level characteristics as explanatory variables for some individual level response variable. This often adopts the common form of a linear-in-means model. Linear-in-means models predict some individual outcome y_i through some composite regressor $\overline{y}$ (or $\overline{y}_{-i}$ if the individual of interest is excluded) that represents the average outcome of the individual's reference group. However, with such models, parameters cannot be correctly identified as first illustrated by Manski (1993). Dubbing it the "reflection" problem, Manski demonstrated that within such models we cannot easily discern between endogenous effects, exogenous effects, and correlated effects when studying how the behavior of a reference group influences individuals of that group.

Given that this project follows in a similar vein of estimating peer effects through utilizing some variant of a linear-in-means model, some discussion on the potential pitfalls of such a model is warranted. As described by Manski, an endogenous effect emerges when an individual's behavior varies with the behavior of the group. This clearly exemplifies endogeneity when an individual's outcome exerts influence on the reference group's outcomes and vice versa. For our purposes, this issue is not of particular concern due to the sequential nature of donations. The donor's reference group is defined as the donations displayed at the time of donation. While these displayed donations may influence an individual's donation, the individual cannot retroactively influence previous donations. Nevertheless, the other effects that Manski describes pose more substantial threats to our analysis. Manski describes exogenous effects as occurring when an individual's behavior varies with exogenous properties of the reference group and correlated effects which appear when individuals in the same reference group act similarly because they have similar individual characteristics. It is easy to imagine how such effects could occur in our context. Donors that contribute close in time to one another may have shared characteristics such as level of income that result in similar observed donation amounts. Yet, with a model that fails to explicitly control for these shared characteristics, the similarity in donation amounts could be taken as evidence of peer effects and not attributed to their true cause. Therefore, throughout this section, we will attempt to mitigate this issue through a number of methods including employing fixed effects and examining the arrival rates of donations. Furthermore, it seems likely that the donations that one observes will not be dictated entirely through events that sequentially target distinct groups of donors. Undoubtably, the donations that one sees will also be subject to some random variation such as when someone sees an announcement or turns on their computer thus creating some sources of exogenous variation.

3.1 OLS Models

Perhaps the most obvious choice for an aggregate value to test for the presence of peer effects is the mean of the displayed donations. Let m be the number of donations displayed on screen. This means that for y_t, the tth donation, the mean value is calculated as $a_t = \frac{1}{m}\sum_{k=1}^{m} y_{t-k}$ where $t > m$. To test for whether this value showed any significance as well as determine whether the effect had any interaction with gender and anonymity, the several linear models were estimated[1]. In addition the mean of displayed donations (histavg), we include a variable to specify whether this was during the period of when five or ten donations were displayed (lagtype) and several indicator variables such as to account for the donor's gender (is_female), whether the donation was made anonymously (is_anonymous), and whether the campaign was a registered nonprofit (is_nonprofit). Since it seems plausible that the number of donations displayed on screen might influence the way a donor responds to displayed donations, we include an interaction term between these variables (histavg:lagtype). Similarly, we also include interaction terms between the mean of visible donations and whether the donation is made anonymously, the gender of the donor, and whether campaign's nonprofit status.

[1]All R source code for various models can be found on my github account, https://github.com/zumiko/thesis. Additionally, residual plots and tests for stationarity are included in the R files

	Model 1	Model 2	Model 3	Model 4
(Intercept)	75.35^{***}	73.09^{***}	97.02^{***}	58.45^{***}
	(2.67)	(2.77)	(1.78)	(3.21)
histavg	0.26^{***}	0.26^{***}	0.18^{***}	0.40^{***}
	(0.01)	(0.01)	(0.01)	(0.02)
lagtypeten	-15.65^{***}	-17.94^{***}	-15.03^{***}	0.73
	(3.19)	(3.22)	(1.72)	(3.65)
histavg:lagtypeten	0.16^{***}	0.19^{***}	0.04^{***}	0.02
	(0.01)	(0.01)	(0.01)	(0.02)
is_anonymous		14.74^{***}		
		(3.45)		
histavg:is_anonymous		-0.06^{***}		
		(0.01)		
is_female			-22.57^{***}	
			(1.65)	
histavg:is_female			-0.13^{***}	
			(0.01)	
is_nonprofit				67.93^{***}
				(8.87)
histavg:is_nonprofit				-0.16^{***}
				(0.02)
R^2	0.03	0.03	0.02	0.03
Adj. R^2	0.03	0.03	0.02	0.03
Num. obs.	234546	234546	161833	233493
RMSE	681.84	681.79	301.52	682.99

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 3.1: OLS regressions for mean of displayed donations

$$\text{Model 1: } \hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1(a_t) + \hat{\beta}_2(\text{lagtype}) + \hat{\beta}_3(a_t * \text{lagtype})$$

$$\text{Model 2: } \hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1(a_t) + \hat{\beta}_2(\text{lagtype}) + \hat{\beta}_3(a_t * \text{lagtype}) + \hat{\beta}_4(\text{is\_anon}) + \hat{\beta}_5(\text{is\_anon} * a_t)$$

$$\text{Model 3: } \hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1(a_t) + \hat{\beta}_2(\text{lagtype}) + \hat{\beta}_3(a_t * \text{lagtype}) + \hat{\beta}_4(\text{is\_female}) + \hat{\beta}_5(\text{is\_female} * a_t)$$

$$\text{Model 4: } \hat{y}_t = \hat{\beta}_0 + \hat{\beta}_1(a_t) + \hat{\beta}_2(\text{lagtype}) + \hat{\beta}_3(a_t * \text{lagtype}) + \hat{\beta}_4(\text{is\_nonprofit}) + \hat{\beta}_5(\text{is\_nonprofit} * a_t)$$

From the Table 3.1[2], we see that the mean of displayed donations appears significant across all of the models. For every one dollar increases the mean we anticipate between around a quarter cent increase in the subsequent donation with all else equal. From Model 1, we see that having ten displayed donations opposed to five appears

[2] I used the R package texreg to produce all regression output tables by Leifeld (2013). .

to increase the size of the effect of the mean of displayed donations[3], though overall donations are lower for this subset of the data. In Model 2, we see that anonymous donations while generally higher, have a signifiant interaction term with the mean anchor value. This negative interaction term implies that anonymous donations are potentially less responsive to the mean of displayed donations than donations where a person chooses to display their name. In Model 3, we also see that donations coded as female tend to be lower than male donations and more notably have a significant interaction with the displayed donations, demonstrating that women are far less influenced by the values of the displayed donations. For a one dollar increase in the mean of displayed donations, a male donor is predicted to give an additional 18 cents as oppose to only 5 cents for a female donor[4]. Finally in Model 4, we see that campaigns with nonprofit status receive donations about seventy dollars higher and are generally less driven by the mean of displayed donations[5].

The R^2 values are quite low ranging between .02 and .03, meaning that these models have very low explanatory power, accounting for only around 2-3% of the variation in the amount donated. However, further inspection reveals that the low R^2 values are be driven down by the subset of the data when ten lags were displayed; the subset of donations with five lags displayed has an $R^2 = .07$ compared to an $R^2 = .03$ for the subset with ten displayed. This suggests that the donations onscreen may be a stronger predictor when there are fewer present, while with more visible reference points people may be less responsive to the reference points, weakening the model. Nevertheless, predictors appear significant suggesting that the explanatory variables do have a discernible effect on the amount donated. Although this may seem somewhat contradictory, it is not a surprising result given that there are a host of omitted variables that would undoubtably exert considerable influence on how much a person might donate such as their level of income and relation to the recipient. Unfortunately, it is difficult to compare this model to the similar work because R^2 values are not included in comparable studies.

As discussed in the last chapter, it is also possible that donations amounts might

[3]Nonetheless, we shouldn't draw the conclusion that having more displayed donations results in a larger peer effects. We might imagine that in this scenario the mean of displayed donations is only serving as a proxy for what type of campaign it is since different campaigns receive different average donations. Therefore, when there are fewer displayed donations the mean would have greater variance and thus be a poorer indicator of the type of campaign.

[4]Note that in this model I dropped observations where the gender could not be identified or the donation was made anonymously which accounts for the lower sample size.

[5]Here there is also a slight decline in the number of observations because due to changes in the way of labeling nonprofits a few campaigns could not be properly identified.

be correlated with factors such as time of day. Such correlations would make it appear as though displayed donations influenced future donations while in reality only detecting the way that donation amounts fluctuate with time of day or season. To account for this effect, additional regressions that included dummy variables for the day of week, hour of day (0-23), and month were considered. The inclusion of these variables was not particularly insightful, the R^2 improved very slightly, but the coefficient on the mean predictor was not diminished. This shows that the effect of displayed donations is probably not attributable to omitted time-based donation patterns. The output of a regression that includes time based dummy variables is included in the first appendix A.2.

We might also wonder whether the scale of campaigns would be a significant determinant in the strength of peer effects. We might expect in a campaign with thousands of donors people are less likely to know one another than in a campaign with only a handful of donors. Therefore, scale of campaign may be able to roughly proxy for familiarity between donors, though of course this relies on a large and somewhat dubious assumption. Still, to see if there was any support for this theory we also ran a regression with the total number of donors as a regressor. While the number of donors itself was not significant in changing the amounts donated, its interaction with the mean of displayed donations was. For every additional hundred donors to a campaign, the effect of the displayed donations is predicted to be two cents less. While this may seem trivially small, it means that for a campaign with 100 donors versus 500, the effect of displayed donations in the former case would be predicted to be 8 cents higher on the dollar which is a pretty notable difference. These results can also be seen in the first appendix A.1.

An additional concern would be that as donations approach some goal amount, donations either increase or decrease in size. If the effect was large, then we would not expect the data to be stationary, however, we will still check if there is a trend apparent. This is related to the work of previously discussed on seed money in Section 2.3 to study the phenomena described by List & Lucking-Reiley (2002) where more seed money leads to both more donations and donations of greater size. To further explore this, we'll calculate the seed money for some donation y_{it}, the tth donation to the ith campaign as $s_{it} = \sum_{k=1}^{t-1} y_{ik}$ (this amount would also be visible to the donor along with the displayed donations). Then we calculate percent raised as $\frac{s_{it}}{g_i} * 100$ where g_i is the goal amount for campaign i. Additionally, we include an indicator variable named goalmet that indicates whether the goal amount has been reached since it seems possible that the effect would be significantly different after that point.

These regression results can be seen in Table 3.2[6]. We see that the mean of displayed donations is still significant with a sizable coefficient. Furthermore, there is no sizable interaction between the amount previously raised and the strength of peer effects. The percent raised variable has a negative coefficient indicating that as the campaign nears its goal amount donations decrease in size. This could be viewed as evidence of previous donations crowding out future donations. However, oddly, once the goal is met, the donations increase as the campaign exceeds the goal amount. Yet, overall donations are lower once the goal amount has been reached. This is a somewhat counterintuitive finding and I think may largely be attributable to the often arbitrary nature of goal amounts on the website; organizers are required to set a goal amount, however they can also change it at any point yet we only have access to knowing the goal amount at the time the donations were scraped. Moreover, many campaigns lack a clear goal amount such as campaigns raising money for a general cause such as funding research into a specific illness. Therefore, I think identifying which campaigns actually had meaningful and constant goal amounts would be important for correctly identifying the effects of a goal amount. Notably, the story with seed money is quite distinct[7]. Additional seed money, such as one more average sized previous donation is expected to marginally increase the subsequent donation. Nevertheless, once the goal amount is reached, additional money makes future donations decrease far more rapidly.

[6]The lower number of observations is due to the fact that once campaigns are terminated by the organizer they no longer display their original goal amount. Additionally, the first 15 donors are excluded from all campaigns.

[7]In this regression, the variable for seed money was scaled by dividing it by 100 around the size of the average donation since this is how much we would expect it to increase with an additional donation and a one dollar increase in the amount raised would not have an sizable effect.

	Model 1	Model 2
(Intercept)	70.21***	36.04***
	(2.46)	(1.66)
histavg	0.38***	0.67***
	(0.00)	(0.01)
percentraised	−0.14*	
	(0.06)	
goalmet	−25.40***	16.75*
	(6.19)	(7.41)
percentraised:goalmet	0.14*	
	(0.06)	
percentraised:histavg	0.00	
	(0.00)	
seedmoney		0.03***
		(0.00)
seedmoney:goalmet		−0.27***
		(0.04)
seedmoney:histavg		−0.00***
		(0.00)
R^2	0.05	0.06
Adj. R^2	0.05	0.06
Num. obs.	179032	185815
RMSE	602.77	590.55

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 3.2: Effect of additional seed money and campaign goal amount

3.1.1 Size variant models

While the previous models are highly indicative of the existence of peer effects, they do relatively little to explain how people decide to donate based on their knowledge of others' donations. Rather, basing prediction off of the mean implies that all displayed values are weighted evenly, or taken into consideration independent of their size, name, gender, and order. In attempt to increase our understanding of which donations are the most influential, we'll now take into consideration the size of the displayed donations.

Looking at the donations according to their relative size can help us to gain insight into whether relatively high or low donations exert more control over a future donation. Allow five new variables d_1, d_2, d_3, d_4, d_5 so that for the jth donation, d_1 would correspond to the smallest donation displayed and $d_1 \leq d_2 \leq d_3 \leq d_4 \leq d_5$. This permutation allows us to see which donations, relatively smaller or larger, are most important in determining the size of the subsequent donation. The results for the linear regression with these new variables is displayed in Table 3.3[8]. The size of the coefficients on this model are significantly larger than in the previous order based regressions. In particular, the large coefficient for the smallest donation seems to indicate a tendency to not want to be the lowest donation on screen. A one dollar increase in the lowest donation results in over a two dollar increase in the prediction for the next donation. Perhaps, the most confusing element is the negative coefficient on the second and third smallest displayed donation, though this can be attributed to the model specification and of course including any one of the predictors does result in a positive coefficient, with decreasing magnitude for larger displayed donations. This model does however carry an interesting implication — by raising the minimum donation displayed onscreen we increase the subsequent donation far more than if the highest donation onscreen increased by the same amount. Another interesting result is looking at the models that only include the minimum, median, or max of the displayed donations. In particular, we observe that the size of the coefficients descend from the minimum to the median to the maximum and that the minimum and maximum donation have greater explanatory power than the median donation.

While we've attempted to understand how the relative size of displayed donations affects subsequent donations, we can also approach this question by examining the effect of uncharacteristically high and low donations, sometimes referred to respectively as "shining knights" and "widows' mites". With the classical model of giving,

[8]This model is just using the subset of data where five donations were present on screen

	All displayed	Min	Median	Max
(Intercept)	37.34^{***}	32.58^{***}	84.47^{***}	78.94^{***}
	(1.59)	(1.66)	(1.50)	(1.45)
d_1	2.41^{***}	2.54^{***}		
	(0.04)	(0.03)		
d_2	-0.43^{***}			
	(0.02)			
d_3	-0.45^{***}		0.26^{***}	
	(0.01)		(0.01)	
d_4	0.30^{***}			
	(0.01)			
d_5	0.05^{***}			0.09^{***}
	(0.00)			(0.00)
R^2	0.17	0.09	0.04	0.09
Adj. R^2	0.17	0.09	0.04	0.09
Num. obs.	68887	68887	68887	68887
RMSE	355.17	373.21	383.82	373.06

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 3.3: Relative size of displayed donations regression

we might expect that large donations could crowd out other giving as discussed in the first chapter. Alternatively, with the private model, large donations could pull donations upwards if people also donate to some extent with social motivations. On the other hand, with small donations it's possible that people may be able to donate less while avoiding social stigma (Smith, Windmeijer, & Wright, 2014).

In Figures 3.1,3.2,3.3,3.4, we identify large donations as being twice the size of the mean donation to that campaign, $y_{it} \geq 2\overline{y}_i$ and small donations as half the size of the mean to that campaign, $y_{it} \leq \frac{1}{2}\overline{y}_i$. This of course, allows for some donations that are coded as large to be smaller than some other donations that are coded as small. To account for that all donations are demeaned, $\ddot{y}_{it} = y_{it} - \overline{y}_i$. Then we display the means of the donations preceding and following, a small donation (Figure 3.1), a large donation (Figure 3.2), and a donation that is neither small nor large (Figure 3.4).

The most apparent feature of Figure 3.1, is that all of the donations on either side of small donations average out to be far beneath the mean. The obvious implication here is that low donations tend to arrive clustered together. This may hint toward the prevalence of correlated effects, as discussed at the beginning of the chapter, or it may support the idea that people base their donations off the donations that are displayed to them. There does not appear to be a clear effect of a small donation and using a Welch's t-test (n = 195857, df = 353563, p = 0.54) on the sum of the ten donations preceding and following the small donation, meaning we cannot reject that the means are the same before ($\mu = -225.13$) and after ($\mu = -210.74$) the small donation. At the same time, the distribution does not look entirely symmetric on either side of the small donation, but that is mainly speculation.

In Figure 3.2, however, there is more evidence of some effect for a large donation. It appears as though donations that occur after a large donation tend to be higher. However, using standard levels of confidence, from a Welch's t-test (n = 39712, df = 79022, p = 0.12) we cannot reject the null in favor of the alternative hypothesis that the mean of the ten donation afterwards ($\mu = 36.30$) is greater than the mean of the preceding ten donations($\mu = 11.22$). Curiously, while the distribution of donations surrounding a small donation looks similar for the subset of the data where only five donations are displayed, with large donations we see a fairly different portrait in Figure 3.3. Here, the mean for the ten prior donations ($\mu = 89.12$) compared to the mean for ten following ($\mu = 163.18$) does prove more statistically significant (n = 4290, df = 7737.4, p = 0.06). While this graph is obviously more turbulent given the smaller sample size, it also encourages us to consider the possibility that high donations might be particularly more impactful when there are fewer donations displayed while the effect is not particularly different for small donations. If this were the case, it could potentially help explain the change in the website layout.

Another general interesting point of comparison, is overall negative mass around a small donation compared to far lesser positive mass around a large donation. This may point to the idea that more people are influenced by the lower donations than people are influenced by higher donations, but perhaps we are moving again to the realm of speculation, since it should also be noted that our criteria for large donations appears harsher than for small donations.

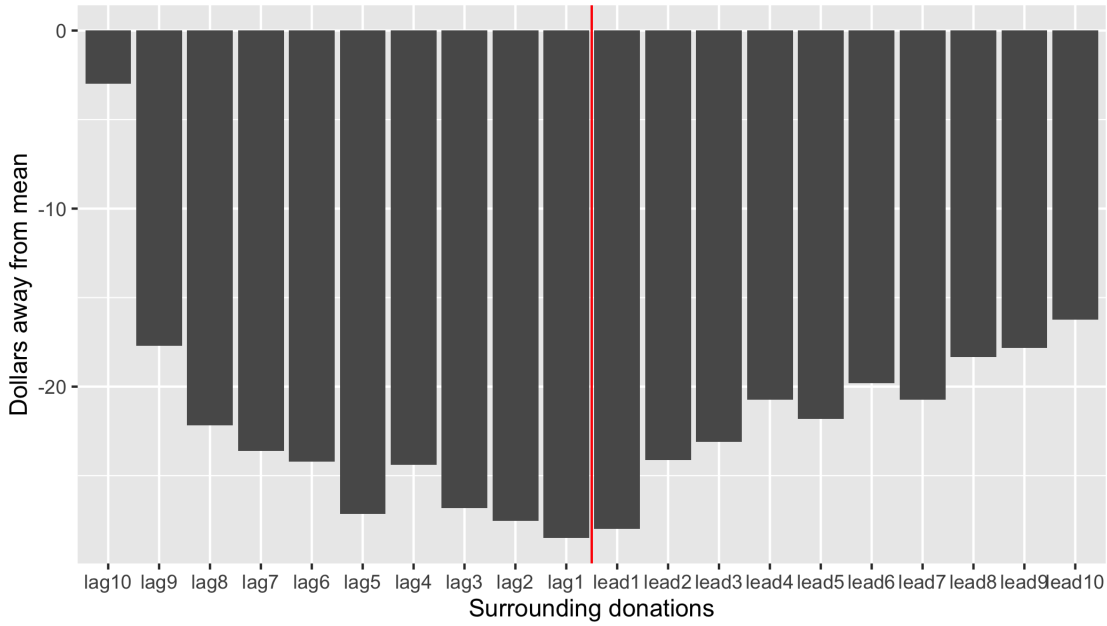

Figure 3.1: Donations prior to and after a small donation

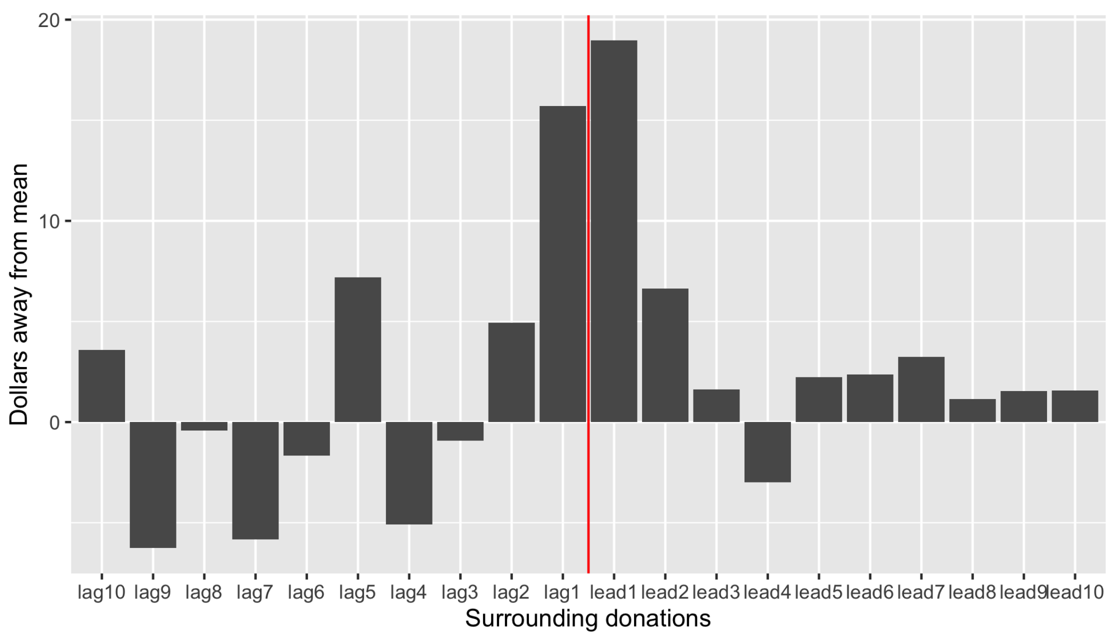

Figure 3.2: Donations prior to and after a large donation

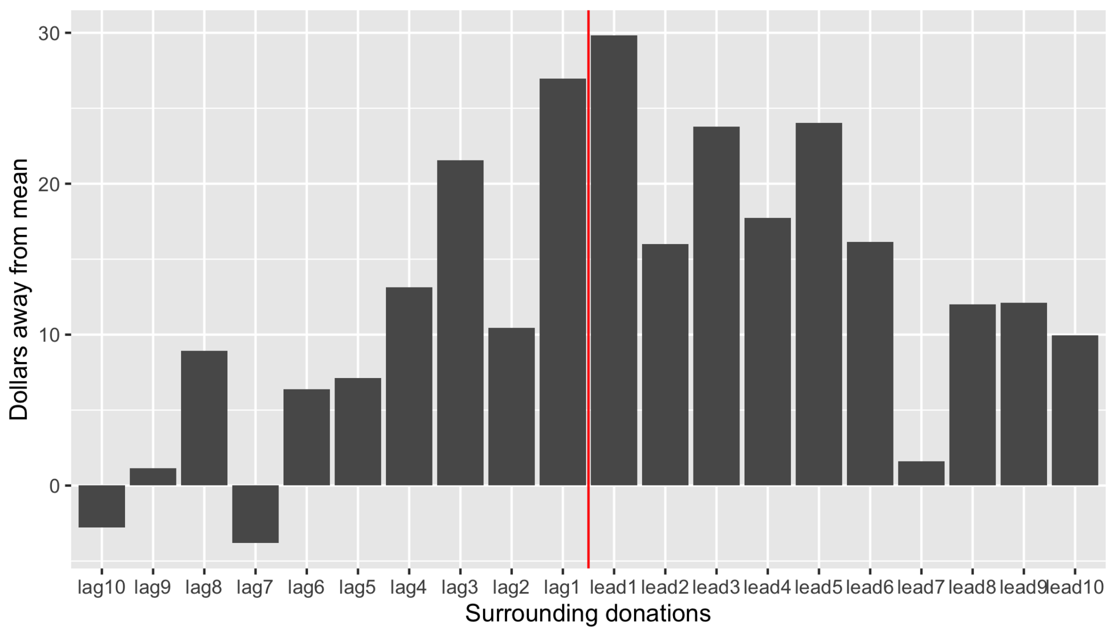

Figure 3.3: Donations prior to and after a large donation (5 displayed)

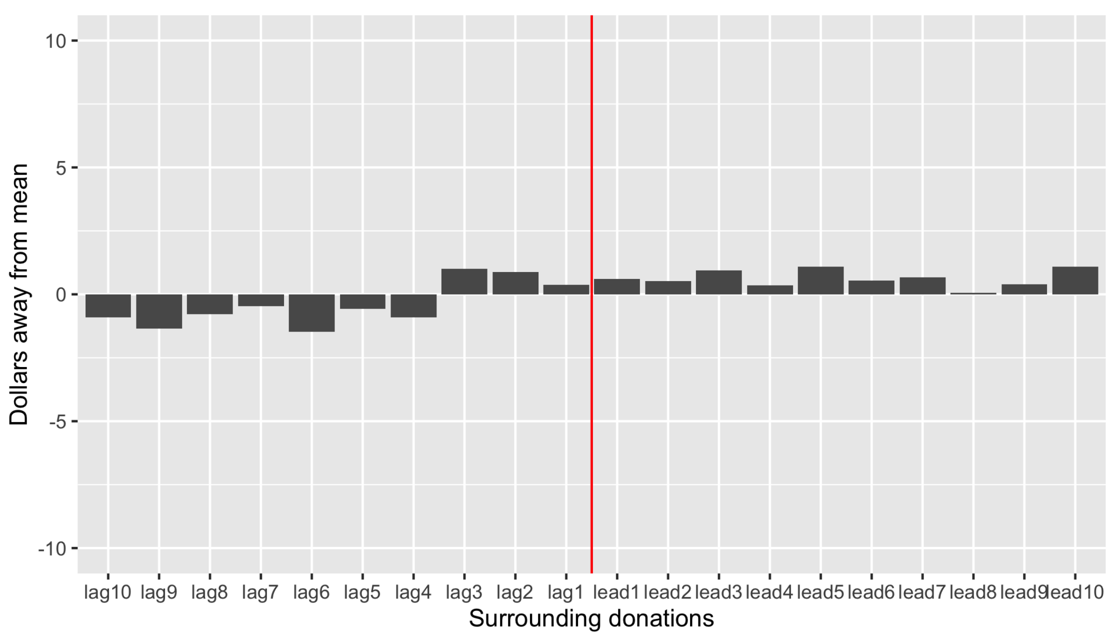

Figure 3.4: Donations prior to and after donations that are neither small nor large

3.2 Quantile Regression

While the previous regressions give us information about how the mean donation responds to the effect of the displayed donations, they tells us little about the overall distribution of the response variable condicional on our predictor variables. To better understand this we now employ quartile regression. There are a number of advantages to this approach. In particular, the least squares estimator is extremely sensitive to the presence of outliers making it a poor estimator for non normal, long tailed distributions (Koenker & Bassett, 1978a). Given the substantial positive skew of donation values this is a clear disadvantage of the previous models. Moreover, there appears to be some heteroscedasticity in the previous models, meaning that while the estimators are still unbiased, OLS is no longer the best linear unbiased estimator. Furthermore, examining the conditional quartiles also provides us with interesting insight into how the displayed donations affect donors at different percentile levels.

In standard OLS estimation we seek to minimize the sum of squared errors, however in the case of quantile regression we instead seek to minimize a weighted sum of the positive and negative error terms so that some proportion of the data τ lies below the regression line and the rest 1 - τ lies above the regression line. That is, as expressed in Koenker & Bassett (1978b), for the linear conditional mean function $E(Y|X = x) = x'\beta$ we find $\hat{\beta}$ by finding β which minimizes, $\sum (y_i - x_i'\beta)^2$. While to find the linear conditional quartile function, $Q_Y(\tau|X = x) = x_i'\beta(\tau)$, we find β that minimizes $\sum \rho_\tau(y_i - x_i'\beta)$ where $\rho_\tau(\mu) = \mu(\tau - I(\mu < 0))$ and I serves as an indicator function. Therefore in the case where $\tau = \frac{1}{2}$ this minimization problem is just reducing the sum of the absolute values of the residuals.

We can see the results of this in Figure 3.5, at three different quantile levels. It is clear from this visualization that at higher quantile levels we observe a much larger effect of the mean of displayed donations. Examining an even greater range of τ values (all tick marks on the horizontal axis), we see a clear trend in Figure 3.6. This holds the potential implication that donors who give high amounts are predicted to increase their donations by more than other donors when the displayed donations on screen increases. This may support the idea that prestige is a particularly strong motive for high amount donors as oppose to other groups of donors. Nevertheless, we should take caution in extrapolating too much from this result given that it is not utilizing the panel aspect of our data as in the former OLS model.

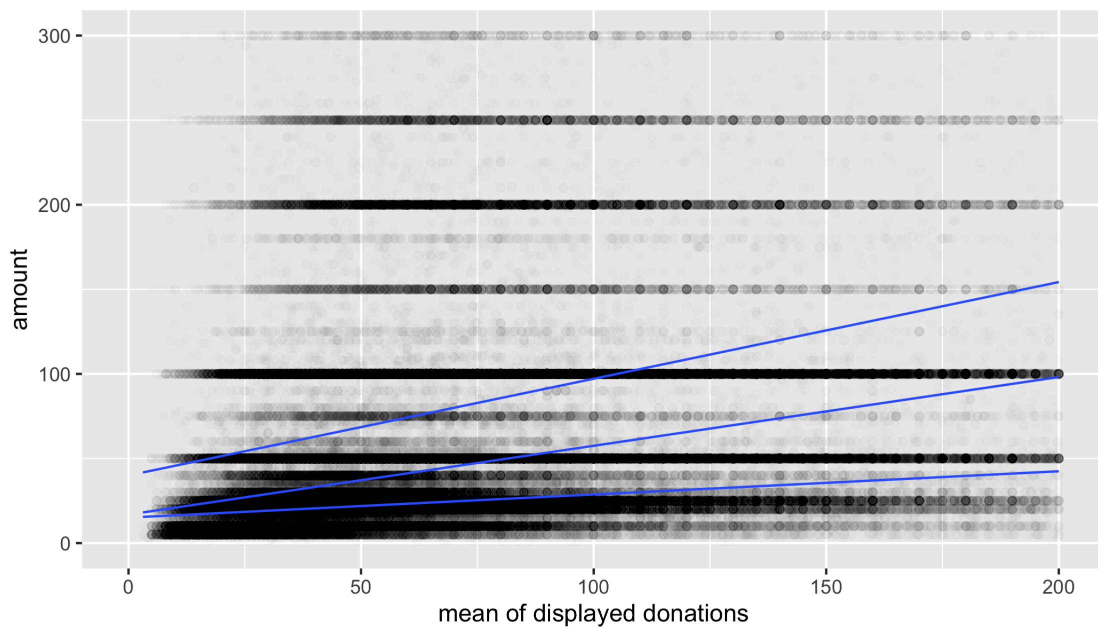

Figure 3.5: Quantile regression lines for $\tau = 0.25, 0.5, 0.75$

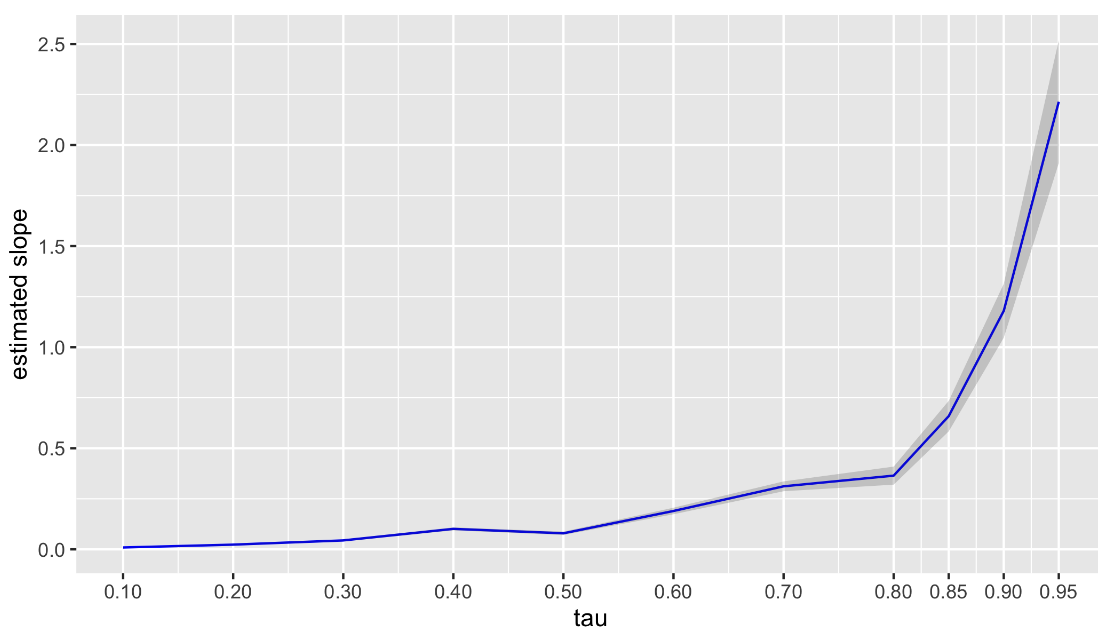

Figure 3.6: Change in slope estimate across quantiles

3.3 Fixed Effects Models

An issue with the previous regressions is that we are not accounting for the differences between campaigns. It seems clear that campaigns will differ substantially from one another. Some may have detailed descriptions and many pictures while others may be relatively sparse. Additionally, each campaign will have a different cause, be located in a different place, have a different recipient, and be part of a different community. We might expect that these features will be correlated with our explanatory variables. For example, a campaign from a tight knit community might have greater peer effects than a campaign for a well-known nonprofit or certain campaigns might attract larger donations from women than from men. Such campaign specific features could result in the detection of biased coefficients because these features are correlated with the parameters we are trying to estimate. A fixed effects model can help resolve this by tersing out time invariate unobserved characteristics of the different campaigns that are likely to be correlated with other explanatory variables.

Let c_i be all the features of the campaign that we do not observe and that are constant across time for some campaign i. Here, the x's are our predictors such as the mean of displayed donations. We expect that c_i may be correlated with many of the explanatory variables that we are using in our model. Therefore, we remove the campaign specific term c_i through time demeaning.

$$y_{it} = \sum_{j=1}^{n} \beta_j x_{itj} + c_i + \mu_{it}$$

Taking the mean across time for each campaign, i,

$$\overline{y}_i = \sum_{j=1}^{n} \beta_j \overline{x}_{ij} + c_i + \overline{\mu_i}$$

Time demeaning the original data,

$$y_{it} - \overline{y}_i = \sum_{j=1}^{n} \beta_j (x_{it} - \overline{x}_i)_j + \mu_{it} - \overline{\mu_i}$$

Relabeling demeaned terms,

$$\ddot{y}_t = \sum_{j=1}^{n} \beta_j \ddot{x}_{tj} + \ddot{\mu}_t$$

	Model 1	Model 2	Model 3
histavg	0.21^{***}	0.15^{***}	0.21^{***}
	(0.00)	(0.01)	(0.00)
is_female		-24.37^{***}	
		(2.59)	
histavg:is_female		-0.13^{***}	
		(0.01)	
is_anonymous			-11.61^{**}
			(4.04)
histavg:is_anonymous			0.26^{***}
			(0.02)
Num. obs.	68895	49043	68895
R^2 (full model)	0.15	0.13	0.15
R^2 (proj model)	0.05	0.01	0.06

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 3.4: Campaign fixed effects (5 displayed)

	Model 1	Model 2	Model 3
histavg	0.09^{***}	-0.03^{***}	0.16^{***}
	(0.01)	(0.01)	(0.01)
is_female		-30.23^{***}	
		(2.18)	
histavg:is_female		-0.01	
		(0.01)	
is_anonymous			28.46^{***}
			(4.82)
histavg:is_anonymous			-0.18^{***}
			(0.01)
Num. obs.	165651	112790	165651
R^2 (full model)	0.06	0.09	0.06
R^2 (proj model)	0.00	0.00	0.00

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 3.5: Campaign fixed effects (10 displayed)

	All displayed	Min	Median	Max
d_1	2.13***	2.27***		
	(0.04)	(0.03)		
d_2	−0.18***			
	(0.02)			
d_3	−0.49***		0.18***	
	(0.01)		(0.01)	
d_4	0.30***			
	(0.01)			
d_5	0.03***			0.07***
	(0.00)			(0.00)
Num. obs.	68887	68887	68887	68887
R^2 (full model)	0.22	0.16	0.12	0.15
R^2 (proj model)	0.13	0.07	0.02	0.05

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 3.6: Fixed effects size based regression

As with the OLS models, we see that while the mean of displayed donations appears significant though the explanatory power of the model is weak. We can see that when there are ten lags of donations displayed that the mean of displayed donations, is notably smaller as seen in Table 3.5 compared to Table 3.4. This raises the possibility that peer effects may act quite differently in the scenario when ten donations are displayed as opposed to when five donations are displayed, creating a broader range of displayed values. In Model 2, we see that the effect of the mean of past donations is far smaller for female donors compared to male donors and even has an inverse correlation with ten displayed donations. Note, however, that when we include all groups of donors, the mean of displayed donations still has a positive effect. Model 3 gives perhaps the most perplexing results. In Table 3.4, we see that anonymous donations appear to exhibit stronger peer effects than other donations, but in Table 3.5 we see that anonymous donations are inversely correlated with the mean of displayed donations. It's possible that using more data could help resolve these contrary findings, but since the two subsets of data are taken several years apart it could even be possible that the social conventions surrounding anonymous giving changed.

Furthermore, we can see that the fixed effect models gains greater explanatory power when all displayed donations are included in the regression are not aggregated as seen in 3.6. It is particularly interesting that the minimum donations holds almost the same explanatory power as the average of donations on screen. Overall, however, these results mirror those found in 3.3 and will therefore not be further discussed.

3.3.1 Arrival Rates

As discussed in the preamble to this chapter, we should be concerned that our data generating process may result in sorting donors into similar groups. A news article or someone sharing the campaign could alert a subset of donors to a certain campaign who would be predisposed to contribute similar amounts. Such events would cause donors with shared characteristics that may affect how much they contribute to arrive in clusters. This seems plausible given that the readership of a particular news outlet or the social network of someone who shares the campaign may be likely to have somewhat similar income levels or values. Omitting such similarities, that are likely correlated with how much people contribute, would cause us to estimate an exaggerated effect. However, such a mechanism, prompting select people to a specific campaign, would likely result in donations that arrive closer together in time. Strong evidence that donations clustered together, or arriving in high frequency, display strong dependence on the displayed donations while the effect is far more limited for more spread out donations would point to endogenous sorting as being a problematic component of our analysis. Therefore, we now consider the interarrival times between the donations.

Let s_{it} be the arrival time of the tth donation to the ith campaign, then we calculate interarrival times by $r_{it} = s_{it} - s_{i(t-1)}$. This is the time that passed between the last donation and the current donation. Then, for each campaign, we divide r_t into quartiles, with smaller interarrival times falling into the first quartile and larger into the fourth quartile. The results of this regression are seen in Table 3.7[9]. The most telling aspect of this table is that the coefficients on the interaction terms for the second and third quartiles are positive indicating that the effect of the displayed donations is greater in size for the subset of donations that arrive with the more time between then compared to the donations that arrive with the greatest frequency. Although there is a negative interaction with the fourth quartile, the size of the interaction is small, and even these furthest spaced donations still demonstrate a large effect of the mean of displayed donations. Overall, there is no clear trend apparent across the quartiles demonstrating that with further spaced donations peer

[9]The number of donations is reduced because I only considered campaigns that had more than 100 donations given that donations are split into quartiles for each campaign. Additionally, the subset of donations with five displayed lags was used because it seems more plausible in this case that displayed donations may all have originated from similar donors.

	Model 1
histavg	0.24^{***}
	(0.02)
quartile2	5.45
	(6.42)
quartile3	-29.34^{***}
	(6.44)
quartile4	28.71^{***}
	(6.38)
histavg:quartile2	0.06^{*}
	(0.03)
histavg:quartile3	0.43^{***}
	(0.03)
histavg:quartile)4	-0.05^{*}
	(0.03)
Num. obs.	38049
R^2 (full model)	0.13
R^2 (proj model)	0.04
Adj. R^2 (full model)	0.12
Adj. R^2 (proj model)	0.03

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 3.7: Interarrival quartiles to test for correlated effects

effects are no longer observable. This makes a tentative case that the effect we detect cannot be attributed to shocks, such as a news article, that drive like minded donors to contribute in clusters.

Conclusion

Throughout this thesis, we've sought to better understand if and how people respond to knowing the donations of others. For our first approach we used the mean of displayed donations to check for the presence of peer effects. We found the mean to be positively correlated with subsequent donations. This finding is in agreement with Smith, Windmeijer, & Wright (2014)[10] and more recently Sisco & Weber (2019). Moreover, this finding is robust to controlling for time and campaign fixed effects, though in the latter case the size of the effect is somewhat diminished. Overall, though estimates vary, the size of the effect is surprisingly large, predicting between around a dime and a quarter increase in the next donation for an additional dollar increase in the mean of displayed donations. This means that when the average of the onscreen donations is \$50 versus \$100, we would expect the next donation in the latter case to be around \$5 higher using the lower estimate. Moreover, we can imagine that this has a type of compounding effect, since by increasing the subsequent donation, the mean of displayed donations will also increase. Although we recognize the potential of correlated effects to inflate the observed effect by examining arrival rates, we believe the size of correlated effects to be limited. However, a more experimental approach where donors are randomly shown a set of values would eliminate the presence of such correlated effects.

Overall, these findings prove consistent with private models of giving that allow for social comparison to dictate the size of the private benefit the donor receives. In other words, a donor's satisfaction from the act of donating depends in part on how their donation compares to what others have given. Nevertheless, the low explanatory power of these models should cause us some pause before concluding that all donors are greatly concerned with how their donation compares to those they ob-

[10]This is not a perfect comparison since the are using the mean of past donations so that for y_{it} the tth donation to campaign i, the mean of past donations is $\frac{1}{t-1}\sum_{k=1}^{t-1} y_{ik}$ while the mean of displayed for this thesis only considers the last five or ten values. Furthermore they find a negative coefficient on this value for campaign fixed effects which they attribute this to the mean differenced error term being negatively correlated with the the mean differenced lagged dependent variable.

serve. Rather, it is possible that a large portion of donors do disregard the additional information they have concerning the donations of others and a select group of donors either use the displayed donations subconsciously as a reference point or take them into account for prestige related reasons.

Indeed, this theory, that a select number of donors may care about the displayed donations far more than others, is supported most clearly by the quantile regression models. In these models, donors that are disposed to give larger amounts based on the displayed donations increase their donations by far more than their lower giving counterparts when the displayed donations increase. This may indicate that such donors are particularly concerned with appearing more generous relative to other donors. Moreover, this trend appears almost exponential. For higher conditional donor level percentiles, the predicted increase in donation size as a result of the displayed donations grows at a rapidly increasing rate. This is a development in understanding the differentiated responses of donors to knowing what others have given and pursuing this line of inquiry is likely worthwhile. In particular, adopting this method to account for the panel aspect of the data would be helpful in assessing the validity of this result. Additional background characteristics of donors, aside from extrapolated gender, such as age and income could also be useful in understanding differentiated donor responses.

Studying the effect of displayed donations across gender, we see that the effect of visible donations seems to exert far more influence on men than women. This may lend support to the idea that men are more concerned with interpersonal comparison than women. Similarly, anonymous donations appear less likely to conform to the displayed donations in our first model. This could imply that people are more likely to donate anonymously if they donate amounts that are less consistent with what has already been donated. However, our fixed effects model pulls this finding into question. Perhaps using the full data collected could help resolve this dispute since knowing how anonymous donations differ could yield insight into the underlying motivations of interpersonal comparison. That is, people may modify their behavior based on others' donations because they personally feel better about their donation when it is relatively higher or they may want for people to see them making a relatively high donation as with conspicuous consumption. By knowing the true effect of anonymity, we could untangle these two distinct motivations, which respectively represent comparison based adaptations of "warm glow" and "prestige" benefits. Overall, this adds an additional facet to our understanding of the differentiated response of donors; the size of the effect of displayed donations seems particularly prominent among those who are inclined to donate in relatively large sums based on the displayed donations and the effect is greatly diminished with female donors.

After testing for the presence of peer effects through the mean value, we then proceeded to try and understand the way in which people took the displayed values into consideration. In particular we looked at large donations and small donations. While it was not clear that large donations increased subsequent donations at standard levels of significance the mean donations following large donations was much higher than preceding them. Furthermore, we saw that low donations appeared to show up surrounded by similarly low donations on either side. Although we can't deny this might be partially attributable to incoming waves of similar donors, it may also indicate that people are only willing to donate in low amounts when others are also donating low amounts.

Another point of curiosity was whether the distribution of the displayed donations was important in determining subsequent donations. To investigate this point, we constructed a series of summary statistics about the displayed donations aside from the mean, such as their kurtosis, skewness, and variance and attempted to gauge whether these distributional statistics would enable us to better predict the amounts donated which can be found in the second appendix. While these models only marginally improved the prediction of donation amounts, this line of inquiry would be an interesting one to pursue further. For example, one could unpack how the effect of displayed donations is different when the donations have high variance relative to low variance. Given that people do exhibit concern for how their donation is situated relative to others' it would seem likely that the distribution of displayed donations would be important if the utility from donating is partially rank-based.

Our size based models may provide some insight into rank based utility. These models indicated that the smallest displayed donation was an equally good predictor as the mean of displayed donations. Additionally, the presence of the large coefficient on the minimum donation, predicting around a two dollar increase in a subsequent donation for a one dollar increase in the minimum, could indicate an unwillingness to donate less than the lowest donation present onscreen. Alternatively, an increase in the maximum displayed donation only results in a very slight increase in the predicted donation amount. This may suggest that what is often referred to as the prestige motive, may have more to do with not being the lowest donation onscreen than about competing to be the highest donation for the majority of donors.

Overall, this thesis may hold some practical results for those interested in increasing fundraising totals. For example, we confirm the previous literature result on the effect of seed money in increasing subsequent donations. Additionally, having larger displayed donations, especially the minimum donation, might heighten future donations, but to a lesser extent with female donors[11]. However, this thesis also attempts to contribute to the thought surrounding charitable giving and interpersonal comparison. In line with private models of giving, there does seem to be a benefit that the donor receives from the act of giving; donations do not shrink as more money accumulates. Moreover, this donor specific benefit seems somewhat contingent on the donations of others. That is, models of warm glow and prestige should account for the idea that such sensations may not enter the donor's utility function directly, but rather as an amount relative to the donations of others. In this way, this project supports the idea that visible charitable donations have a clear social component, meaning that a donor's willingness to contribute is neither fixed nor independent, but instead shaped by their knowledge of what others have given.

[11] However, I think displaying the top donations would not have the same effect. Rather, I think the most recent donations may cause people to believe that the values they see are indicative of what is typically given creating a distinct effect.

Works Cited

Andreoni, J. (1989). Giving with impure altruism: Applications to charity and ricardian equivalence. *Journal of Political Economy*, *97*(6), 1447–1458. https://doi.org/10.1086/261662

Andreoni, J. (2006). Leadership giving in charitable fund-raising. *Journal of Public Economic Theory*, *8*(1), 1–22. https://doi.org/10.1111/j.1467-9779.2006.00250.x

Bagwell, L. S., & Bernheim, B. D. (1996). Veblen effects in a theory of conspicuous consumption. *The American Economic Review*, *86*(3), 349–373. http://www.jstor.org/stable/2118201

Darwin, C. (1871). *The descent of man: and selection in relation to sex*. London: J. Murray.

Frank, R. H. (1985). The Demand for Unobservable and Other Nonpositional Goods. *The American Economic Review*, *75*(1), 101–116. Publisher: American Economic Association. https://www.jstor.org/stable/1812706

Friedman, D., & Ostrov, D. N. (2008). Conspicuous consumption dynamics. *Games and Economic Behavior*, *64*(1), 121–145. https://linkinghub.elsevier.com/retrieve/pii/S0899825608000092

Glazer, A., & Konrad, K. A. (2008). A signaling explanation for charity. In *40 Years of Research on Rent Seeking 2*, (pp. 713–722). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-79247-5_43

Greenwell, B., Boehmke, B., Cunningham, J., & Developers, G. (2019). *gbm: Generalized Boosted Regression Models*. R package version 2.1.5. https://CRAN.R-project.org/package=gbm

Harbaugh, W. T. (1998). The prestige motive for making charitable transfers. *The American Economic Review*, *88*(2), 277–282. http://www.jstor.org/stable/116933

Heller, N. (2019). The hidden cost of gofundme healthcare. *The New Yorker*. https://www.newyorker.com/magazine/2019/07/01/the-perverse-logic-of-gofundme-health-care

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An Introduction to Statistical Learning*. Springer New York. https://doi.org/10.1007/978-1-4614-7138-7

Koenker, R., & Bassett, G. (1978a). Regression Quantiles. *Econometrica*, *46*(1), 33. https://www.jstor.org/stable/1913643?origin=crossref

Koenker, R., & Bassett, G. (1978b). Regression Quantiles. *Econometrica*, *46*(1), 33. https://www.jstor.org/stable/1913643?origin=crossref

Leifeld, P. (2013). texreg: Conversion of statistical model output in R to LaTeX and HTML tables. *Journal of Statistical Software*, *55*(8), 1–24. http://www.jstatsoft.org/v55/i08/

Liaw, A., & Wiener, M. (2002). Classification and regression by randomforest. *R News*, *2*(3), 18–22. https://CRAN.R-project.org/doc/Rnews/

List, J. A. (2011). The market for charitable giving. *Journal of Economic Perspectives*, *25*(2), 157–180. https://doi.org/10.1257/jep.25.2.157

List, J. A., & Lucking-Reiley, D. (2002). The effects of seed money and refunds on charitable giving: Experimental evidence from a university capital campaign. *Journal of Political Economy*, *110*(1), 215–233. https://doi.org/10.1086/324392

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The Review of Economic Studies*, *60*(3), 531. https://doi.org/10.2307/2298123

Mason, R. (2000). The social significance of consumption: James duesenberry's contribution to consumer theory. *Journal of Economic Issues*, *34*(3), 553–572.

McCormick, K. (1983). Duesenberry and veblen: The demonstration effect revisited. *Journal of Economic Issues*, *17*(4), 1125–1129. http://www.jstor.org/stable/4225388

Moffitt, R. (2000). Ipolicy interventions, low-level equilibria and social interactions. *Social Dynamics,*.

Monroe, R. (2019). When gofundme gets ugly. *The Atlantic*. https://www.theatlantic.com/magazine/archive/2019/11/gofundme-nation/598369/

Mullen, L. (2019). *gender: Predict Gender from Names Using Historical Data*. R package version 0.5.3. https://github.com/ropensci/gender

Rablen, M. D. (2008). Relativity, Rank and The Utility of Income. *The Economic Journal*, *118*(528), 801–821. https://academic.oup.com/ej/article/118/528/801-821/5089514

Romano, R., & Yildirim, H. (2001). Why charities announce donations: a positive perspective. *Journal of Public Economics*, *81*(3), 423–447. https://doi.org/10.1016/s0047-2727(00)00078-5

Sisco, M. R., & Weber, E. U. (2019). Examining charitable giving in real-world online donations. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-11852-z

Smith, S., Windmeijer, F., & Wright, E. (2014). Peer effects in charitable giving: Evidence from the (running) field. *The Economic Journal*, *125*(585), 1053–1071. https://doi.org/10.1111/ecoj.12114

Stewart-Williams, S. (2015). Morality: Evolution of. In *International Encyclopedia of the Social & Behavioral Sciences*, (pp. 811–818). Elsevier. https://linkinghub.elsevier.com/retrieve/pii/B9780080970868810569

Tullock, G. (1966). Information without profit. *Public Choice*, *1*(1), 141–159. https://doi.org/10.1007/bf01718991

Veblen, T. (1899). *The Theory of the Leisure Class*. Macmillan.

Vesterlund, L. (2006). Why do people give. *The Nonprofit Sector: A Research Handbook*.